Digitized by the Internet Archive
in 2015

https://archive.org/details/differentialeffe00sher

Department of Psychology

THE DIFFERENTIAL EFFECTS OF ASSESSMENT CONTEXT AND SCORING METHOD ON CREATIVITY PERFORMANCE IN CHILDREN

by

David William Sherwood

Department of Psychology
Duke University

Date: Oct 17, 1968

Approved:

Donald J. Stedman
Donald J. Stedman, Chairman

Lloyd J. Borstelmann

Michael A. Wallach

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Psychology in the Graduate
School of Arts and Sciences of Duke University

1968

20

ABSTRACT

(Psychology--Clinical)

THE DIFFERENTIAL EFFECTS OF ASSESSMENT CONTEXT AND SCORING METHOD ON CREATIVITY PERFORMANCE IN CHILDREN

by

David William Sherwood

Department of Psychology
Duke University

Date: ________________

Approved:

____
Donald J. Stedman, Chairman

____

____

____

____

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Psychology in the Graduate School of Arts
and Sciences of Duke University

1968

ABSTRACT

THE DIFFERENTIAL EFFECTS OF ASSESSMENT CONTEXT AND SCORING METHOD ON CREATIVITY PERFORMANCE IN CHILDREN

by

David William Sherwood

Of two child creativity assessment approaches evolving from the work of Guilford, that of Wallach & Kogan has identified a coherent, IQ-independent creativity dimension, while that of Torrance has not. The present investigation examined the possibilities that differences in assessment context and scoring method were responsible for these discrepant findings. Forty white, sixth grade males, heterogeneous with respect to IQ and SES, individually were administered five representative Torrance, and three representative Wallach-Kogan creativity procedures under evaluative (EV), test-like conditions. An additional forty Ss received the same tasks under nonevaluative (NEV), game-like conditions.

All procedures were scored for *fluency* (number of appropriate responses); *originality* (number of statistically infrequent responses); and *flexibility* (number of different conceptual categories into which responses fit). In the analyses of corre-

lational patterns, scoring method was varied by (a) considering all three sets of scores together, and (b) analyzing flexibility scores separately from fluency and originality scores. It was predicted that the presence of either EV conditions or the flexibility score would increase correlations between creativity and intelligence variables, and reduce correlations among the creativity variables.

It was found that differences in assessment context as defined in this study did not significantly affect creativity-intelligence correlations. Under either EV or NEV conditions, the Wallach-Kogan measures intercorrelated in the +.50's ($p<.01$), while the four Torrance Verbal measures intercorrelated in the +.20's ($p<.10$). With the exception of the single Torrance Figural task, no creativity procedures shared significant variance in common with IQ. Mean creativity scores were generally higher under NEV conditions, but not significantly so. The flexibility score did not relate to IQ more than did fluency and originality, nor did it decrease the average intercorrelations among creativity variables. Flexibility thus seemed appropriate as a creativity measure. No social class differences were found with respect to mean creativity scores or perception of assessment context. Finally, while "high IQ" Ss (mean IQ=117) did not differ from "low IQ" Ss (mean IQ=96.4) on mean creativity scores, high IQ Ss tended to score higher in EV conditions on certain tasks, while low IQ Ss scored higher in NEV conditions on the same tasks.

It was concluded that there exist differences between Torrance and Wallach-Kogan procedures which are not attributable

primarily to differences in scoring method or assessment context as defined in this study. It was suggested that future research examine the effects of group *vs.* individual, and timed *vs.* untimed assessment conditions.

ACKNOWLEDGMENTS

The doctoral dissertation, like almost any "independent" project, seldom is completed without the cooperation and direct assistance of a great many people. I wish to express appreciation to the principals and teachers who gave their time and resources in order to make data collection possible. Principals of participating schools were: Mr. Fenner Boyd, Mr. Herman Nelson, Mr. Baxter Ridenhour, Mr. Edwin Warehime, and Mr. Willis Whichard. The teachers of the eight classes involved were: Mrs. Virginia Atkinson, Mrs. Lana Brunner, Mr. Floyd Eamon, Mrs. Helen Elliott, Miss Mary Fox, Mrs. Gail Von Frank, Mrs. Nancy Waldrep, and Mrs. Celeste Watkins. Thanks are also extended to Mrs. Florence McCracken, of the Durham City School System, who made initial contacts with participating schools.

Particular appreciation is due to Dr. Robert Spaulding and Dr. Donald Stedman for arranging financial and technical assistance for this research through the Durham Education Improvement Program, sponsored by the Ford Foundation.

Many people provided helpful feedback, opinions, and criticisms of various aspects of the research. Thanks are due to the members of the Psychology Department at the Astor Home for Children in Rhinebeck, New York--particularly to Drs.

Elliott Brown, Michael Kulla, and Donald Johnson, for serving as judges in the scoring reliability check. Many helpful suggestions were made by the members of my doctoral committee: Drs. Irving Alexander, Lloyd Borstelmann, Alan Kerckhoff, Darwyn Linder, Donald Stedman, and Michael Wallach. I am indebted particularly to Dr. Stedman for serving as Chairman of my committee, and for providing consistent support, motivation, and guidance throughout the planning and execution of the study.

To my wife, Paula, goes my gratitude and appreciation not only for her invaluable assistance as typist and critic, but more importantly, for her support and understanding in meeting all the challenges which a dissertation inevitably entails.

D.W.S.

CONTENTS

ABSTRACT iii

ACKNOWLEDGEMENTS vi

LIST OF TABLES x

I. INTRODUCTION 2

Background, 2
The Empirical Research, 8
Goals of the Present Investigation, 24

II. METHOD 29

Subject Characteristics, 29
General Procedure, 33
The Creativity Procedures, 38
Scoring, 41
Principal Analyses, 47

III. RESULTS 51

Subject Characteristics by Assessment Condition, 51
Effects of Assessment Context, 52
Effects of Scoring Method, 61
Effects of Test Source, 64
Effects of Socioeconomic Status, 66
Effects of Intelligence, 67

IV. DISCUSSION AND SUMMARY 70

Relation of Findings to Previous Research, 70
Implications for Further Research, 85
Summary, 88

APPENDIXES 93

A. HOLLINGSHEAD'S TWO-FACTOR INDEX OF SOCIAL POSITION (MODIFIED) 93

B. CREATIVITY TASK INSTRUCTIONS 95

EV Condition, 95
NEV Condition, 99

C. SCORING GUIDE FOR JUDGING "SAMENESS-DIFFERENCE" IN WALLACH-KOGAN RESPONSES 103

D. TABLES 104

REFERENCES 117

LIST OF TABLES

(See Appendix D)

1. Guilford's Divergent Thinking Factors
2. Comparison of EV and NEV Treatment Groups on Five Selection Variables
3. Intercorrelations among the Nine Wallach-Kogan Creativity Measures and with Intelligence for EV (N=40)
4. Intercorrelations among the Twelve Torrance Verbal Creativity Measures and with Intelligence for EV (N=40)
5. Intercorrelations between the Torrance Verbal and Wallach-Kogan Creativity Measures for EV (N=40)
6. Intercorrelations between the Creativity Indexes, Intelligence, and Total Number, Originality, and Flexibility Scores for EV (N=40)
7. Intercorrelations among the Nine Wallach-Kogan Creativity Measures and with Intelligence for NEV (N=40)
8. Intercorrelations among the Twelve Torrance Verbal Creativity Measures and with Intelligence for NEV (N=40)
9. Intercorrelations between the Torrance Verbal and Wallach-Kogan Creativity Measures for NEV (N=40)
10. Summary of Creativity-Intelligence Relationships for All Number, Originality, and Flexibility Variables for EV (N=40)
11. Summary of Creativity-Intelligence Relationships for All Number, Originality, and Flexibility Variables for NEV (N=40)
12. Intercorrelations between the Creativity Indexes, Intelligence, and Total Number, Originality, and Flexibility Scores for NEV (N=40)

13. Intercorrelations between "Repeated Figures" Variables and Other Creativity and Intelligence Variables for EV and NEV Conditions

14. Results of One-way Analyses of Variance across Assessment Condition for All Measures of All Procedures, and Totals

15. Summary of Creativity-Intelligence Relationships for Number and Originality Variables for EV (N=40)

16. Summary of Creativity-Intelligence Relationships for Flexibility Variables for EV (N=40)

17. Summary of Creativity-Intelligence Relationships for Number and Originality Variables for NEV (N=40)

18. Summary of Creativity-Intelligence Relationships for Flexibility Variables for NEV (N=40)

19. Intercorrelations between Modified Creativity Indexes, Intelligence, and Total Number and Originality Scores for EV and NEV

20. Summary of Creativity-Intelligence Relationships by Assessment Condition and Scoring Method

21. Results of Two-way Analyses of Variance between Social Class and Assessment Condition for All Measures of All Procedures

22. Results of Two-way Analyses of Variance between Intelligence and Assessment Condition for All Measures of All Procedures

THE DIFFERENTIAL EFFECTS OF ASSESSMENT CONTEXT AND SCORING METHOD ON CREATIVITY PERFORMANCE IN CHILDREN

CHAPTER I.

INTRODUCTION

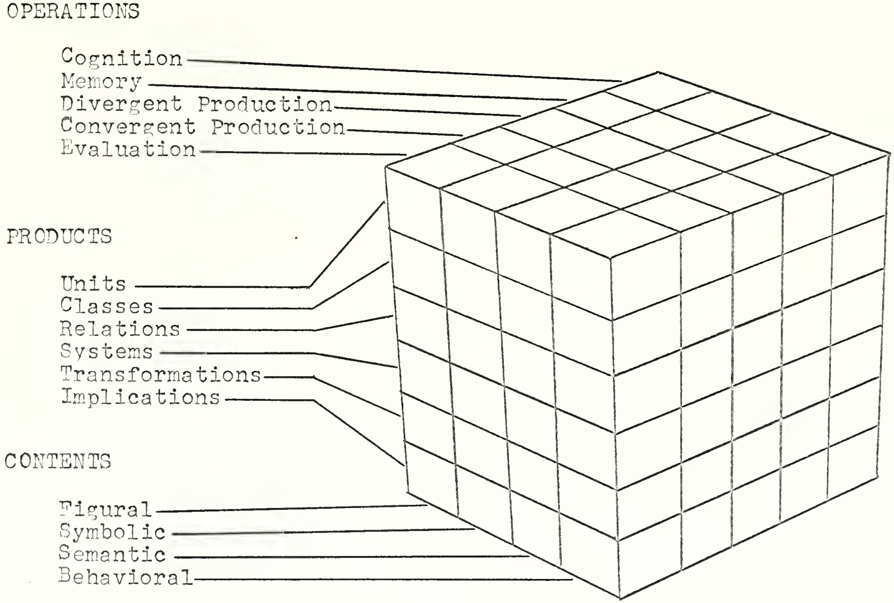
Background

In the measurement of individual differences, recent developments have included consideration of creativity as a group of mental abilities conceptually distinct from, but analogous to intelligence. Measures have been devised to tap so-called basic "creative" abilities in the same way that "IQ tests" measure traditionally defined intellectual abilities. Empirical findings differ, however, with respect to the degree of statistical relationship among creativity measures and between measures of creative and intellectual abilities. The present investigation developed out of a need to examine, and perhaps to reconcile, conflicting findings arising from two major creativity assessment approaches with children. It has as a further aim the refinement of existing research methodologies to more adequately and efficiently tap creative abilities which are relatively separable from intelligence.

The Guilford model. The major pioneer effort in developing objective measurement of creative abilities has come from the Guilford cognitive model, the "Structure of Intellect." This work is presented here as a theoretical basis for the creativity-intelligence distinction, and as a background for the more recent creativity assessment programs to be compared in the present investigation. (See Guilford, 1956, 1957, 1959, 1965, 1966, for the evolution of the "Structure of Intellect.") This model, developed from a factor-analytic approach, attempts to classify all mental abilities within a system defined by the intersections of values on three dimensions. These dimensions include the kind of operation or process involved, the type of product which results, and the content category of any given mental activity. As of the 1966 revision, 80 out of the 120 postulated abilities have actually been measured. A diagram of the "Structure of Intellect" model is presented in Figure 1. (From Guilford, 1966, p. 21)

Of special relevance are two of the five subcategories in the "Operations" dimension, entitled "Convergent Production" (or "Convergent Thinking"), and "Divergent Production" (or "Divergent Thinking"). The essence of convergent thinking involves the channeling of thought toward a single right answer--the thought process typically measured by traditional intelligence tests. On the other hand, Guilford says of divergent thinking factors: ". . . within whatever limits that are imposed by external restrictions, the need for rejecting or

Figure 1. The "Structure of Intellect" Model



superseding a response and for trying or producing a new one is the common element in this group of factors." (Guilford, 1956, p. 281) That is, rather than converging toward a single right answer, divergent thinking involves the search for many suitable responses in many directions. While Guilford does not limit creative abilities to this realm, divergent thinking represents the first major attempt to operationalize and distinguish creativity from intelligence. At present, Guilford claims to have isolated eight divergent thinking abilities: four involving some type of *fluency* of word or idea production; three involving *flexibility* in conceptual activity; and one involving *originality* (novelty or uniqueness) of productions. These eight factors, along with brief examples of measuring tasks, are summarized in Table 1.

Table 1. Guilford's Divergent Thinking Factors

Factor	Description and Examples
1. Word Fluency	The ability to produce words according to structural requirements, regardless of meaning (e.g., anagrams).
2. Associational Fluency	The ability to generate words to fill a specific category of meaning (e.g., word association tasks).
3. Ideational Fluency	The ability to generate ideas meeting specified conditions (e.g., "Name all objects which are round and edible.").
4. Expressional Fluency	The ability to transform ideas into words (e.g., "Make a sentence for the frame, 'W___ c___ e___ n___.'").

Table 1. (continued)

Factor	Description and Examples
5. Adaptive Flexibility	The ability to transcend a common set (e.g., matchstick problems).
6. Spontaneous Flexibility	The ability to shift conceptual categories freely while generating ideas; freedom from perseveration.
7. Redefinition	The ability to shift from old interpretations of familiar objects in order to perceive a new or different function (e.g., seeing that a fishbone can become a needle).
8. Originality	The ability to give responses that are judged clever, or are statistically unusual or unique (e.g., making up plot titles).

Sources: Wilson, Guilford, Christensen, & Lewis, 1954; Guilford, 1956, 1959, 1965.

These abilities may be summarized by suggesting that they all involve a capacity to transcend fixed, ordinary channels of thinking--to "move out on one's own," as it were. It is this independence from the conventional which appears to be one of the basic conditions for creative thinking. Three of these factors--ideational fluency, spontaneous flexibility, and originality--are of special importance, for they represent the scoring measures most often used to form a composite creativity index in subsequent assessment programs. They are considered in greater detail in the later discussion of the Torrance and Wallach-Kogan research.

The "common variance" approach. Consider again the two major categories from the "Structure of Intellect" model which

are intended to represent approximately the domains of intelligence and creativity--"convergent" and "divergent" thinking. In making a theoretical distinction between these or any two broad concepts, one must present some basis for giving them different labels. Otherwise, a common label would suffice. Thorndike (1963) illustrates this point well:

> When we use any term to designate an attribute of an individual, be it abstract intelligence, sociability, or creativity, the distinct label implies that there exists a set of behaviors of the individual that can be grouped together because they exhibit a common quality. Furthermore, if the label is a useful one, the set of behaviors that it designates can be differentiated from the sets designated by other labels . . . We may appropriately ask how well the attribute 'creativity' meets these joint criteria of designating a reasonably extensive set of behaviors which (1) have some degree of coherence and (2) can be distinguished from other sets of behaviors. (Thorndike, 1963, pp.44-45)

Guilford (1959), Torrance (1966), and others would argue that the diversity of creative abilities overrides their commonality, and hence it does the term "creativity" an injustice to be considered a unitary dimension distinct from intelligence. Wallach (in press), however, argues cogently that the approach of looking for significant variance shared in common among divergent thinking measures that is not shared with intelligence is the more conservative approach. Such conservatism is necessary because there is no adequate validity criterion for most creativity assessment procedures, especially with children and adolescents. Teacher and peer ratings, quality of children's drawings, and other behavioral measures of creativity have been shown to be largely unacceptable as validity criteria, partially because of their contamination with intelligence factors. (Reid, King, & Wickwire, 1959;

Torrance, 1963; Wallach & Kogan, 1965) The "common variance" approach achieves its validity intrinsically, through the formation of an internally coherent dimension. Wallach asks:

> Can assessors of divergent and convergent thinking be distinguished in terms of differences that are large enough to be psychologically meaningful in the resultant orderings of the respondents? This is the crucial question that must be asked of the proponents of the structure of intellect model. (Wallach, in press, p. 12, MS.)

Wallach argues that whether one accepts the Spearman "G" or Thurstone "Primary Mental Abilities" concept of intelligence, there exists among specific regions of content a set of uniformly positive and significant correlations accounting for the major portion of variance in intellectual assessors. Similarly, if one wishes to speak of a set of creative abilities as distinct from intelligence, he must likewise demonstrate a set of uniformly positive and significant correlations among specific creativity measures, which share relatively little variance in common with intelligence measures. The more clearly such an independent dimension can be formed, the more justifiably the assessment techniques involved may be said to be measures of creativity or divergent thinking rather than simply inferior measures of intelligence.

The Empirical Research

Of two major creativity assessment approaches evolving from Guilford's theoretical framework and utilizing modifications of his measuring instruments, that of Wallach & Kogan (1965) has empirically demonstrated a coherent creativity dimension,

whereas that of Torrance (1966) has not. Two principal differences between these approaches emerge as possible explanations for their discrepant results. The first difference involves the assessment context; the second involves the scoring method. The present investigation seeks to examine the role of these two factors in both Torrance and Wallach & Kogan assessment procedures, to determine their relative contributions to the formation of a unitary and virtually intelligence-free creativity dimension.

The Torrance program. Both the Torrance and Wallach-Kogan programs are attempts to make Guilford's assessment of divergent thinking appropriate for administration to children; hence they include adaptations of many of his measures. Most of Torrance's published research is based on a battery of creativity measuring tasks which extend beyond Guilford's divergent thinking factors. This battery is described fully by Torrance (1962). More recently he has published a "research edition" (Torrance, 1966) based on refinements of the earlier battery, and including administrative, scoring, and normative manuals for both verbal and figural tasks. Verbal procedures include such tasks as guessing the causes and consequences of an event, and thinking of unusual uses or improvements for an object. Included in the figural tasks are the composition of pictures from various line segments, and the drawing of as many pictures as possible based on a repeated stimulus such as a circle. Scoring categories are fluency (number of responses elicited per task item); flexibility (number of conceptual categories to

which responses belong); originality (number of statistically infrequent or unusually imaginative responses); and optionally, elaboration (number of nonessential details added). The first three scoring measures derive from the Guilford divergent thinking factors discussed earlier (Table 1, pp. 5-6). Scores on individual tasks are summed to give a composite creativity index. The battery is declared suitable from kindergarten through graduate school; however, the content is clearly more appropriate to the elementary age range. Each task is timed rather strictly, and the battery is group administered in the classroom above the third grade level.

The bulk of the available evidence relating the Torrance composite index to measures of intelligence fails to demonstrate an intelligence-free creativity dimension. First Torrance's own work, then findings of other researchers using the Torrance or a similar index, are considered.

Torrance's fundamental approach has been to show that subjects scoring within the top twenty percent on the composite index, but not on intelligence, achieve as well academically as those scoring within the top twenty percent on intelligence, but not on creativity. He argues that such students achieve creatively, rather than intellectually. In eight partial replications of the Getzels & Jackson (1962) study (Torrance, 1962), including elementary through graduate school subjects, six showed no significant differences in school achievement between creative and intelligent groups. This finding becomes open to question, however, when one realizes that the mean IQ of the

creative groups was still far above average. Two groups were graduate students; the mean IQ's of the other four were 126, 122, 122, and 112. Moreover, in the remaining two cases where the high intelligence groups showed higher achievement than the high creativity groups, the latter had substantially lower IQ's (98 and 107). This finding lends further support to the conclusion that in the first six groups, both highly intelligent and highly creative subjects were actually achieving primarily on the basis of superior intelligence.

Another argument frequently used by Torrance (e.g., Torrance, 1962, 1963) is that of these subjects falling within the top twenty percent on the composite creativity index, only about thirty percent also score within the top twenty percent on intelligence. He presents this as evidence for the relative independence of the measured creativity and intelligence domains across a wide age range. This conclusion is questionable on two counts. First, there are no creativity test intercorrelations presented, so the meaning of creativity is ambiguous. Second, and more important, the comparison of two variables, one or both of which is unduly restricted in range (i.e., the "top twenty percent"), has been demonstrated to produce spuriously low intercorrelations. (See Madans, 1964; DeMille & Merrifield, 1962; Ripple & May, 1962.) Hence, it is not unlikely that average correlations between the full range of creativity and intelligence scores in any sample would suggest considerably more variance held in common than is indicated by Torrance's argument.

A recent attempt by Torrance and his colleagues to explain the creativity-intelligence relationship is the "threshold of intelligence" notion, which suggests that creativity and intelligence are uncorrelated only above a critical IQ level of about 120. (Torrance, 1963) If this is so, addition of a creativity index should predict academic achievement above this level significantly better than an IQ measure alone. Wallach (in press) concludes, however, in summarizing several relevant studies (e.g., Yamamoto & Chimbidis, 1966), that while IQ is a less adequate predictor of academic achievement at high than at moderate or low IQ levels, addition of a Torrance creativity index adds almost nothing to predictability of achievement at any level. Hence, he continues, the Torrance index, summed from the four components of fluency, flexibility, originality, and elaboration, behaves like a second measure of convergent thinking, at least with respect to prediction of academic achievement.

Torrance does not utilize the "common variance" approach discussed above to analyze the relationship between his index and intelligence measures; however, such an approach yields conclusions similar to those already reached. For example, Yamamoto (1965) found in two large samples of fifth graders (one middle class, the other heterogeneous with respect to social class), correlations of +.33 and +.39 respectively, between the Torrance index and Lorge-Thorndike Verbal IQ. The similar findings in the two groups suggest also that class differences have little effect on the relationship between creativity and intelligence. An earlier study by Wodtke (1964) found the Torrance

index correlated with the Lorge-Thorndike total IQ at +.44 in a sample of fourth graders, and +.45 in a similar sample of fifth graders. These correlations are all moderate and significant.

Most studies do not provide comparable intercorrelations among Torrance creativity measures, since the measures are typically pooled into the composite index. Circirelli (1964), however, has correlated all pairs of eight creativity measures (fluency, flexibility, originality, and elaboration, for the verbal and figural parts of Torrance's battery), obtaining an average intercorrelation of +.37. This figure is inflated by the intrinsic correlation between fluency and originality scores; it is closer to around +.30 (Wallach, in press), or slightly less than the typical correlation between the Torrance index and intelligence.

Summaries of additional research using the Torrance and similar Guilford indexes have been provided by Ward (1966); Wallach & Kogan (1965); and Wallach (in press). The findings of Ward are typical. Averaging results from seven Guilford studies, he finds a mean correlation among creativity measures of +.20, with a mean correlation between creativity and convergent thinking measures of +.21. He then concludes: "There is no indication, then, that Guilford's divergent thinking factors define a broad, unitary domain, or that they have any variance in common that is not shared with IQ." (Ward, 1966, p. 17) Similarly, Wallach & Kogan (1965) cite a number of studies (e.g., Getzels & Jackson, 1962; Cline, Richards, & Needham, 1963; Flescher, 1963), in which significant average correlations of about +.30

were found to exist both between creativity measures and IQ, and among the creativity measures themselves.

In summary, the measures included in the Torrance composite creativity index (and similar Guilford indexes) appear to share at least as much, if not more, of their variance with intelligence as with each other. If this is so, they become, in effect, inferior measures of intelligence, rather than measures of a distinct creativity domain.

The Wallach-Kogan program. This research (Wallach & Kogan, 1965) was undertaken in an attempt to measure the hypothesized intelligence-free, internally coherent creativity dimension. A total of 151 fifth grade, normal, middle class children were administered five creativity procedures and a number of intellectual and personality tests. Verbal creativity procedures included giving instances of a specified class of objects (e.g., "Name all the round things you can think of."); suggesting alternate uses for common objects such as a newspaper or a button; and proposing similarities between common objects such as a cat and a mouse. The two remaining procedures employed visual stimuli--abstract lines and patterns--for which the child was asked to give as many meanings or interpretations as possible. Procedures were scored for fluency (total number of responses), and uniqueness (occurrence only once in the sample). This scoring system omits the flexibility and elaboration scores of Torrance. Finally, all creativity procedures were administered individually as games, without time limits or known evaluative pressures of any kind.

The resulting internally consistent creativity dimension proved virtually unrelated to intelligence. Intercorrelations among creativity measures (both verbal and visual) ranged from +.07 to +.74, with a mean of +.41. A vast majority of the creativity intercorrelations were significant at the .01 level. At the same time, correlations between creativity and intelligence were low and generally nonsignificant, averaging +.10. These correlations were based on the total range of both creativity and intelligence scores in the sample tested.

The evidence presented thusfar indicates, then, that the Wallach-Kogan creativity measures share both more variance among themselves and less variance with intelligence than do the Torrance measures. To what extent might differences in assessment context and scoring method account for these discrepant patterns?

The effect of assessment context. The issue concerning the psychological climate in which any ability tests are administered is an important one. Everyone performs somewhat differently under different conditions; children are particularly sensitive to environmental cues to facilitate optimum performance. Zigler & Butterfield (1968) have recently challenged the appropriateness of standard intelligence-testing conditions, at least with culturally deprived children. They have shown that measured IQ can increase substantially when testing is done under conditions which maximize success experiences. A similar increase in measured IQ over the course of a nursery school year was likewise shown to be the result of a reduction in the effect of "debilitating motivational factors." Do assessment conditions

similarly affect the amount or pattern of measured creativity in children?

In virtually all of the Torrance and Guilford-based studies, creativity procedures were administered under "evaluative," or test-like conditions. Administration took place in groups under strict time pressure, with instructions to do one's best, just as on a standard intelligence or achievement test. In the Wallach-Kogan assessment, by contrast, every precaution was taken to eliminate the threat of pressure or evaluation by providing individual assessment in a relaxed, game-like framework. Theoretical and empirical evidence suggests that evaluative assessment conditions both inhibit creative production, and inflate the correlation between divergent and convergent thinking measures by eliciting convergent thought processes. Wallach & Kogan offer this explanation:

> . . . thinking of various types may be enhanced or diminished in quality simply by altering the context in which the child's thinking is assessed . . . . Where the assessment context emphasizes success and failure, we suspect that differentiation within the thinking domain will be minimized. Under those conditions, the forms of thinking demanded by conventional intelligence tests can be expected to exercise dominance over a wide span of cognitive tasks. As the assessment situation becomes more permissive, however, the overriding effects of intelligence should decline, with the consequence of greater differentiation in the thinking domain. (Wallach & Kogan, 1965, pp. 305-06)

The importance of a nonevaluative, permissive environment in facilitating creative production is stressed by the "self-actualizing" schools of personality theory. (e.g., Anderson, 1959, 1965; Rogers, 1959, 1963; Maslow, 1959) A more empirically oriented explanation is the "associative hierarchy" theory of Mednick (1962). In Mednick's terms, the creative

person has a less steep gradient of associative strength than does the non-creative person. In producing a stream of associations which fit specified conditions, the creative person, although beginning with stereotyped responses, will continue producing increasingly unique responses long after the non-creative person has exhausted his repertory with the end of his stereotyped responses. Because his gradient is flatter, the creative person responds at a relatively slow rate, although he will give both a higher total and a higher proportion of unique responses. It is evident in this formulation that the creative person's measurable performance will be less than optimal unless conditions permit both temporal freedom and a playful or nonevaluative task context. Temporal freedom allows the creative person to complete his more extensive response repertory; too short time limits may actually make the non-creative individual appear more productive, because of his initially more rapid response rate. That a playful assessment context seems necessary to facilitate the generation of associations is supported by the introspections of creative people such as Einstein, Poincaire, and Housman. (Ghiselin, 1952) These observations point to the necessity of temporary suspension of critical and judgmental functions, which is best achieved under nonevaluative conditions.

Research supports the facilitating role of both the temporal freedom and the nonevaluative qualities of a relaxed, game-like atmosphere. Word association studies shed light on the role of temporal freedom in creative association. Siipola,

Walker, & Kolb (1955), comparing the effects of temporal pressure vs. freedom on simple word association tasks, found that subjects given temporal freedom produced more responses of greater associative distance from the stimulus. That is, a noun stimulus elicited an adjective response, rather than another noun. The authors describe subject reports of rich mental imagery elicited by stimulus words in the freedom condition, which in turn called forth response words conveying the idiosyncratic meaning for the subject. Apparently the extra time allowed personal stimulus variables to affect verbal habits. The authors suggest that under temporal pressure, by contrast, the set for speed produced a selective effect, such that the subject automatically limited his responses to those which required the least associative leap--those that were more stimulus-bound. According to Mednick's formulation, these responses are also the more common or stereotyped. This prediction received support in a study by Horton, Marlowe, & Crowne (1963), in which more "original" associations appeared under temporal freedom than under strict time limits.

Studies involving the personal evaluation aspect point to a similar conclusion. Reece & Whitman (1962), for example, varied the assessment condition in a simple fifteen minute free association task, with no imposed requirements for response appropriateness. In the evaluative condition, E drummed his fingers, leaned away from the subject, and did not smile; in the nonevaluative condition, he appeared accepting and interested. In a simple fluency measure, more free associations were produced

in the nonevaluative condition.

Two other studies used uniqueness, or originality, as the dependent variable. Dentler & Mackler (1964) gave a timed "Uses" task from Guilford under four levels of evaluation, ranging from supportive and "psychologically safe" through detatchment, aloofness, and threat. Three times as many uncommon responses were produced in the "safe" condition as in any of the other three. In an intriguing use of role-playing technique, Wild (1965) had subjects perform word-association and object sorting tasks under two role sets in addition to an initial spontaneous or natural condition. The evaluative set instructed the subject to respond as a "regulated" person, concerned about what others thought of him; the nonevaluative set had the subject respond as a relaxed, carefree person. Instructional sets were based on two character sketches. A greater increase in originality from the initial natural condition occurred in the nonevaluative role set.

The effect of scoring method. Differences in method of scoring may provide a second explanation for the discrepant correlational patterns in the Torrance and Wallach-Kogan programs. It was mentioned earlier that the Torrance creativity index is formed by summing the divergent thinking factors of fluency, flexibility, and originality, with the optional addition of an elaboration score. In order for this composite score to describe usefully an intelligence-free dimension, each of its component scores should be statistically relatively independent of common intelligence measures, as Thorndike (1963) has sug-

gested. Furthermore, its component scores should possess at least as much variance in common as any one does with IQ. If any one of the component scores is more highly correlated with IQ than with the remaining creativity components, its effect is to increase the variance in common between intelligence measures and the composite creativity index, thus reducing the effectiveness of all the components as assessors of the hypothesized creativity domain. The present section examines available evidence independently relating fluency, flexibility, and originality measures to IQ measures, in order to ascertain which ones share the least variance with intelligence.

Bereiter (1961), reanalyzing data from an earlier study by Gewirtz (1948) which compared fluency measures and Stanford-Binet Vocabulary scores on five- and six-year-olds, found a correlation of +.01 between ideational fluency and the IQ measure. Additionally, ideational fluency was found negatively related to a third factor--word fluency (r=-.24), suggesting that this divergent thinking factor might better be considered as falling outside the realm of creativity.

More recently, Orpet & Meyers (1966) approached the question by using a factor analysis with an orthogonal varimax rotation. All divergent thinking procedures were administered individually to a large sample of children. Factor analysis revealed a clear ideational fluency factor, with divergent thinking measures loading from +.57 to +.66. Intelligence measures loaded +.30 and below on this factor. Conversely, on the intelligence factor, where the WISC subtests loaded +.71 and +.69,

ideational fluency loaded from +.02 to +.22. Considerable separation between intelligence and fluency scores is apparent here.

A study by Feldhusen, Denny, & Condon (1965) sheds light on the relationship between intelligence and all three divergent thinking factors in question. Selected Torrance tasks ("Alternate Uses" and "Consequences"), along with intelligence and achievement tests were group-administered under standard testing conditions to subjects in the seventh and eighth grades. The resulting correlation matrix showed fluency unrelated to either IQ (average r=+.05) or achievement (average r=+.04). On the other hand, flexibility was strongly related to both (average r's=+.47 and +.44, respectively). Originality was more ambiguously related, correlating about +.28 with IQ and achievement for boys, and +.11 for girls. By contrast, Clark, Veldman, & Thorpe (1965), using originality scores derived from three Guilford tasks given to junior high students, found an average correlation of +.04 with intelligence measured by the California Test of Mental Maturity.

Wallach's (in press) reanalysis of three Guilford studies found similar patterns among seven of the eight proposed divergent thinking factors. If one deletes "word fluency," which appears mediated by intellectual factors, the remaining three fluency factors show somewhat greater intercorrelation among themselves than with IQ. For example, in the study by Guilford & Christensen (1956), these three fluency factors intercorrelated +.31 among themselves, and +.22 with intelligence. In the Wilson, Guilford, Christensen, & Lewis (1954) study, two fluency

factors correlated +.28 between themselves, and only +.08 with intelligence. Finally, in a study by Guilford, Frick, Christensen, & Merrifield (1957) the lone fluency factor correlated +.01 with intelligence. The three flexibility factors correlated +.13 among themselves and +.15 with intelligence in the Wilson, *et. al.* study; the corresponding figures from Guilford, *et. al.* were +.24 and +.20. Originality was inconsistent, with correlations with intelligence of +.25, +.01, and +.19, respectively in the three studies.

Available evidence from Torrance and Guilford research thus suggests that the composite creativity index derives much of its common variance with IQ from the flexibility component, and to a smaller extent from the originality component. Parallel evidence regarding the elaboration component is unavailable; however, it seems likely that elaboration follows flexibility in being more closely related to intelligence than to remaining creativity measures. Wallach & Kogan's successful isolation of the independent creativity dimension may thus be attributable in part to their restriction of scoring to fluency and originality--the two measures most intrinsically distinct from measures of convergent thinking.

Such a restriction may be necessary because it isolates the nexus of the creative process--the simple ability to generate ideas to a given stimulus. This basic operational definition of creativity minimizes the role of verbal, conceptual, or representational skills mediated by intelligence. Wallach (in press) suggests:

> Maximally independent from convergent thinking ability and coherent in itself when measured in various ways, therefore, may be a person's ability to produce large numbers of basically appropriate ideas in response to a given task, and--at least partly because of generating a more extensive idea flow--his ability also to produce more ideas that are relatively unique or unusual. The other kinds of traits placed by the Guilford group in the "divergent thinking" category, on the other hand, do not seem in terms of the available findings to belong there. (Wallach, in press, p. 42, MS.)

The implication is that flexibility, elaboration, and other factors generally considered basic aspects of creativity may in fact be mediated intellectually and only peripherally related to the basic creative process.* It is granted that most productive and socially recognized creativity relies heavily upon intellectual skills for expression; however, the successful measurement of a clear-cut creativity domain may necessitate "cutting through" such intellectual mediators to the fundamental creative process. This is what Wallach & Kogan appear to have done.

In short, evidence presented suggests that measures of fluency and originality appear relatively independent of intelligence, even under evaluative conditions. Furthermore, recall that fluency and originality appear to be facilitated by relaxed, evaluation-free assessment conditions, with time pressure minimized.# These two findings combined suggest that Wallach & Kogan's success in establishing a coherent and independent

---

* It might be argued, for example, that flexibility, involving shifts among conceptual categories, is more dependent on IQ because of the need to apply rules of logic and concept-formation, which are IQ-related functions. (Wallach, in press)

# The effect of such conditions on flexibility and elaboration scores remains to be seen.

creativity dimension is attributable to both provision of optimal conditions for creative expression, and utilization of measures least intrinsically contaminated by intellectual functions.

Goals of the Present Investigation

The statistical isolation of the creativity dimension by Wallach & Kogan is, they feel, ample demonstration that creativity is best elicited and measured under nonevaluative conditions, scoring only fluency and originality; the failure of Torrance and Guilford researchers to isolate such a dimension is attributed to the absence of these implicitly necessary and sufficient conditions. Thusfar, however, this notion has not been tested by examining the creativity-intelligence correlational patterns of representative Torrance procedures administered and scored identically to those of Wallach & Kogan. If Torrance's procedures could be shown to tap the same unified creativity dimension when administered and scored under Wallach & Kogan's conditions, much support could be provided for the "nexus of creativity" notion. Furthermore, such a finding would greatly increase the usefulness of the Torrance procedures, for they would thereby more efficiently tap the creative process and would be less "contaminated" by intellectual factors. By increasing their validity in this manner, these procedures could become more meaningful in studies correlating creativity findings with other personality variables.

On the other hand, by administering and scoring Wallach &

Kogan's tasks under Torrance's conditions, it should be possible to determine the relative contribution of assessment condition and scoring method to their unified dimension, by noting to what extent this dimension breaks down. If, however, administration and scoring variables have little effect on the correlational patterns associated with the two assessment programs, one would then have to look further for intrinsic differences in their creativity assessors. The primary goal of the present investigation is to answer these major questions and to thereby determine the optimal and most efficient assessment and scoring conditions for measuring basic creativity in children.

Assumptions. The present investigation rests on four basic assumptions:

1. A fundamental aspect of creativity, more specifically defined as the ability to generate ideas--both common and unique, can be measured by paper and pencil tests, given orally and individually.

2. Creativity is operationally distinguishable from intelligence.

3. It is possible to create evaluative (test-like) and relaxed, nonevaluative (game-like) assessment climates through manipulation of external cues such as task instructions and demeanor of the examiner.

4. These two sets of conditions can be created in children of both middle and lower socioeconomic status; lower class children can perceive the game-like climate as relaxed and nonevaluative, as can middle class children.

Hypotheses. Based on the findings discussed in the present chapter, this investigation tests the following predictions:

1. Effects of assessment context.
    a. Test-like assessment conditions appear to elicit intellectual, rather than creative, thought processes; therefore, based on an average of fluency, originality, and flexibility scores, both the Torrance and Wallach-Kogan procedures given under evaluative (EV) conditions will correlate as much or more with IQ as with each other. Average correlations between creativity measures and IQ will be significant.
    b. Relaxed, evaluation-free assessment conditions appear to allow creative thought processes to predominate over intellectual ones; therefore, based on an average of fluency, originality, and flexibility scores, both the Torrance and Wallach-Kogan procedures given under nonevaluative (NEV) conditions will correlate as much or more among themselves and between each other as with IQ. The average of intercorrelations among all creativity measures will be significant.
    c. Relaxed, evaluation-free assessment conditions appear to allow maximum creative expression; therefore, all mean creativity scores will be higher under NEV than under EV conditions, across IQ and SES.

2. Effects of scoring method.

   a. Fluency and originality appear to be enhanced by NEV conditions, while flexibility appears intrinsically related to intelligence; therefore, the correlational patterns predicted in hypotheses 1a and 1b will emerge strongly for fluency and originality together, but will fail to emerge for flexibility alone.

   b. By the same reasoning, mean fluency and originality scores will show a stronger increase from EV to NEV than will mean flexibility scores.

3. Effects of test source.

   There is no clear evidence that the Torrance and Wallach-Kogan procedures are intrinsically different in principle or in ability to measure creativity; therefore, when similar assessment conditions and scoring methods are employed, no significant differences will emerge between Torrance and Wallach-Kogan procedures with respect to (a) intercorrelations among creativity measures, or (b) correlations with IQ.

4. Effects of socioeconomic status.

   a. Lower class children are frequently hampered in verbal facility (Bereiter & Englemann, 1966; Smith, 1965); therefore, middle class Ss will obtain higher creativity scores than will lower class Ss on verbal creativity procedures, regard-

less of assessment context.

b. Assumption 4 may be tested by hypothesizing that the effect of assessment condition on overall creativity scores will be the same for both middle class and lower class Ss.

CHAPTER II.

METHOD

Subject Characteristics

Eighty white male subjects (Ss) were selected from eight sixth grade classes in five Durham, N.C. city schools. They were chosen to represent a cross section of social classes and geographic locations in Durham. The number of Ss provided by each school ranged from four* to 39. Subjects were selected to be homogeneous with respect to race, sex, and age, but heterogeneous with respect to intelligence (IQ) and socioeconomic status (SES).

Race and sex. It was felt that limiting the sample to white males avoided racial and sex differences which would have increased the complexity of the design and made cell sample sizes too small to yield meaningful comparisons. White Ss were chosen over Negroes because previous research using the creativity measures in question has been limited to Caucasian samples, and because there is some evidence (Smith, 1965) to suggest that more

---

* In one school, by school policy, letters requesting parental permission for children's participation were sent out. The four Ss were those whose parents gave their permission, and who met the remaining qualifications for acceptance in this study.

meaningful intra-race differences in measured creativity emerge in white than in Negro Ss.

Males were selected over females for three reasons. First, with the exception of a sex-role-linked task in Torrance's earlier battery (ways to improve a toy fire engine and nurse's kit), neither Torrance nor Wallach & Kogan have found any consistent sex differences in level or pattern of creativity scores at the upper elementary level. (Torrance, 1962, 1966; Wallach & Kogan, 1965) Second, Wallach & Kogan found that males seemed to respond slightly better to nonevaluative assessment conditions than did females. Finally, since a secondary purpose of this investigation was to compare creativity patterns in children from divergent socioeconomic backgrounds, boys were chosen over girls by the reasoning that social class differences are more salient for males because of their role of future family head and provider.

Age. The age level selected (eleven years through thirteen years, five months) is only slightly older than that of the Wallach-Kogan sample, and avoids the developmental depressions in measured creativity described by Torrance (1964) at the fourth and seventh grade levels. Furthermore, children at this age level possess the necessary verbal skills to understand and deal with the task instructions, but they have not yet begun to encounter the pressures of adolescence, with whatever effect such pressures might have on creativity. Finally, in studies involving comparisons between creativity and intelligence, the problem arises that IQ scores are age-corrected, whereas creativity

scores generally are not. This problem is easily controlled by restricting the sample to a relatively narrow age range.

Intelligence. The intelligence measure used was the group-administered *Lorge-Thorndike Intelligence Test*, Level 3, Form A, suitable for grades four through six. Both verbal and nonverbal batteries were administered. A highly reliable, valid, and well-standardized instrument, the *Lorge-Thorndike* has a standard error of 4.4 IQ points, and correlates +.77 with the WISC. (Lorge & Thorndike, 1954) The standardization sample of 136,000 children in 22 states represents the normal proportions of various socioeconomic levels found in the overall population. Therefore, this instrument is particularly well-suited to a sample heterogeneous with respect to socioeconomic status. A total IQ score was computed by taking the simple mean of the verbal and nonverbal IQ scores, and rounding upward when necessary. To be included in the study, an S had to score 80 or above on either the verbal or nonverbal portions, and on the total IQ score. It was felt that an IQ of 80 represented the minimum intellectual level needed to comprehend and carry out adequately the creativity task instructions. On this basis a total of three in the eight classes tested were rejected. No further restrictions were placed on IQ, as it was desired to tap a wide intellectual range representative of the normal population. The total IQ in the present sample ranged from 84 to 133, with a mean IQ of 106.

Socioeconomic status. Participating schools were selected to represent both middle and lower class neighborhoods; two out of the five schools included Ss from every category. Socio-

economic status was measured by Hollingshead's Two Factor Index of Social Position (1957), which provides a weighted rating on the basis of father's occupation and educational level. (See Appendix A)

Weighted ratings are classified into five groups, I to V, with I generally representing major professionals and higher executives with at least a four-year college degree, and V representing unskilled laborers or public welfare recipients with a minimum of schooling. Information concerning father's occupation and education, as well as age of S, was obtained from school cumulative folders or from teachers when school records were vague or incomplete.* Children for whom scoreable information was not obtainable were dropped from the study. Twenty-one Ss were eliminated in this way.

One problem encountered was whether to include children from broken homes in the sample, as it was unknown how their atypical background might affect their creativity. While it was necessary to exclude several of these children due to insufficient background information, it was decided to retain the remaining seven in order to represent more accurately the normal population. These Ss were classified according to the occupation and education level of their guardians. Only one of these Ss came from class I; the remainder fell into classes IV and V.

---

* Several teachers obtained information on father's occupation by assigning class themes on "What My Father Does at Work." The children were not informed of any connection between these themes and the present study.

In summary, then, a subject was selected for the present study if he was male, white, in the sixth grade, scored above 80 on either verbal or nonverbal and total IQ, and could be classified into one of five socioeconomic levels on the basis of father's or guardian's occupation and education. Selection continued until an N of 80 was obtained.

General Procedure

Initial contacts were made with the principals of all the Durham city elementary schools. Of those able to cooperate, five were selected as representative of the Durham population. The study was described to the principals and involved teachers as an attempt to compare creativity and intelligence scores under different assessment conditions. Teachers were requested to refrain from any mention of "creativity" or "testing" in their classes in reference to the examiner.

Intellectual assessment with the Lorge-Thorndike was carried out at the beginning of the study, during the second week in November, by another examiner. It was decided to test all children in each of the eight classes in order to avoid questions and confusion which might have resulted from testing only the white boys. Testing was done in the classrooms, and was presented as part of the regular school assessment program.

During the same week, the writer examined all potential Ss' school records to obtain age and SES information. When his presence at a school coincided with that of the intelligence

examiner, great care was taken not to be seen by the children and possibly connected with the intellectual testing.

Creativity assessment took place from the middle of November through January. The general plan was to complete the assessment in one school or class before starting the next one. Upon beginning with a class, the examiner introduced himself and his purposes to the entire class in the following way:

> I'm very interested in children your age, and I'd like to talk with some of you individually during the next few days. Let me explain briefly what we'll be doing. Actually, I'm interested in two different things. Some of you will be taking some new kinds of tests to see how you do. With others, there won't be any tests at all. Instead, I'd like to play some new games with you in order to find out what kinds of games different children like best. I can't tell you ahead of time which group you'll be in, but the groups are decided by flipping a coin. So if your friend comes back and says he took the tests, you might get to play the games instead. I wish I could talk to everybody, but since I have so many other children to see at other schools, I'll be talking with just boys, and just some of the boys. After I've talked with you, please don't tell the others who haven't gone yet what you did, so they can be surprised too. Okay, let's start with (name of child).*

Children were then taken one at a time in random order from the classroom. Total creativity assessment time varied widely among individual Ss, ranging from 45 minutes to over

---

* Originally it was hoped to be less specific about the creativity procedures, to avoid either the "test" or "game" set prior to administration. It was also felt that probably the best way to prevent the children from discussing the procedures among themselves was to say nothing about secrecy. However, lack of the above explanation in the first class tested led to excessive curiosity, which the teacher felt was disruptive, and which apparently prompted some of the later Ss in this class to ask their tested classmates what they had been through. Careful questioning of these later Ss prior to assessment revealed that, in two or three cases, they were expecting to be tested and had a very rough idea of one or two of the measures; however, after careful assurance to the two in the NEV condition that they

two hours. Average time was about one and a half hours. Because of school schedules involving breaks for lunch and physical education, no child remained in the assessment situation for more than an hour and a half at one time. This appeared well within the attention and interest span of most Ss. In the few cases where boredom or fatigue became apparent, the session was terminated and resumed at a later time during the same day.

The facility for testing was adequate, but less than ideal in many schools. Testing rooms were usually makeshift, ranging from an auditorium stage to a small book storage room. A desk or table and two chairs were available in all rooms. In those rooms containing numerous objects which could have been distracting, S was seated so that he faced minimum distraction.

Ss were divided randomly into evaluative (EV) and non-evaluative (NEV) assessment conditions, with approximately half the Ss from each class falling into either group. In the overall sample, exactly half, or forty, Ss were placed in each condition.

Evaluative condition. Under the EV condition, the examiner (E) was friendly, but businesslike, minimizing irrelevant conversation. Upon seating himself opposite S, he proceeded:

---

were not being tested, their prior knowledge and set did not appear to affect their attitude or responses. The distinction between "tests" and "games," and the injunction not to discuss the procedures, made at the suggestion of the first teacher, appeared very effective in limiting curiosity and discussion among children, and controlling initial set to the experimental conditions.

In this case, Ss in neither EV nor NEV conditions can be considered to have been "tested" in the traditional sense, because test scores were not used in relation to their school work or life situation.

> You have been put in the group that is to take the tests. I have some new kinds of tests here that you've probably never seen before. They're sort of like intelligence tests--you know, the kind where you have to pick the right answer out of the four or five that are given. Well, instead of just one right answer, these tests have many right answers. There are also many wrong answers, but the idea is to get as many of the right ones as you can. These tests can tell us a lot about how good you are at thinking in different ways. They will be scored and your scores will be compared with those of other children just as in any test, so it's very important to do the best you can. I will be timing you with this stopwatch. Do you have any questions? (Answer questions in general terms.) Then let's begin with this one. (See Appendix B for EV instructions for specific tasks.)

A stopwatch was placed in S's view, and total time elapsed was recorded for each of the eight tasks. The overall purpose was to create a moderately stressful atmosphere typical of the standard intelligence or achievement test, with implied time pressure and a premium placed on "correctness" of responses. Many comments made by Ss in this condition suggested that the desired task set was probably achieved. The eight creativity measures were then administered in a preselected random order different for each S.

Nonevaluative condition. The procedure with Ss in the nonevaluative (NEV) condition began with a warmup period designed to maximize rapport and to put S at ease. After introducing himself as someone interested in playing some new and interesting games with children, E handed S a pencil, paper, and crayons, suggesting:

> Before we try some games, I'd like to give you a chance to draw a picture of whatever you'd like. This is just to relax and get into a game-playing mood, so don't worry about how good it looks. Just draw whatever you want to.

During this period, E offered informal comments designed to support and relax S, and develop rapport. When the drawing was complete, S was told he could keep it for himself.

Then E proceeded:

> I'm interested in finding out what kinds of games different children like. I have quite a few new games here that you may not have seen before, but other children have enjoyed playing them. I'd like to try some of these with you just to see which ones *you* like the best. There are different kinds of games--some are word games, some are picture games, and there is also a drawing game. Most of the games are played by your telling me things just as they pop into your head. Since this is just for fun, there's no such thing as a right or wrong answer. The things we say and do together don't really count for anything--except to help me see which games different children like the most. So, since everything is just between you and me, don't be afraid to say whatever comes into your head. Do you have any questions about what we're going to do? (Answer questions in general terms.) Okay, let's begin with a game I think you'll like. (See Appendix B for NEV instructions for specific tasks.)

No stopwatch was used, and all hints of time pressure were avoided. Every effort was made to maintain a relaxed, friendly rapport with S and to displace any hint of evaluation from S to the games themselves. To follow up this way of presenting the tasks, each S was asked at the end of the session which "games" he enjoyed playing most, and which he least enjoyed. Their responses indicated that nearly all Ss in this condition enjoyed the overall session, and apparently did perceive the measures as games. As in the EV condition, each S received a different predetermined random order of administration, to control for order effects.

In both conditions, responses were given orally, accepted without comment, and recorded by E on standard answer forms.* Where a response was unintelligible or ambiguous, however, S was asked to repeat or elaborate his response, e.g., "Can you tell me

---

* The one exception to this is Repeated Figures, a drawing task, described below.

more about that?"

In order to prevent sessions from becoming excessively lengthy, yet to preserve temporal freedom, the following system was devised. It was used under both assessment conditions, in order to keep roughly equivalent the amount of time each S was allowed to spend at a given task. Nondirective encouragement was given to S's responses in the form of comments like "fine," "okay," "mm-hmm," etc. and prompting questions such as "Anything else?" or "Those are fine; did you have any other ideas?" The nondirective comments were used during short pauses; the prompting questions were asked when approximately ten seconds or more had elapsed since the last response or prompting. No more than three prompting questions were asked during any specific task. The task was ended when approximately fifteen seconds had elapsed following the third attempt at encouragement,* or when S indicated with some finality that he was through.#

The Creativity Procedures

To keep the duration of testing moderate and interest optimal, four procedures from the Torrance Verbal, one from the Torrance Figural, and three from the Wallach-Kogan batteries were

---

* In the NEV condition, this timing was done inconspicuously with E's wristwatch.

# If S in either condition said he was through with a task before at least one encouragement had been given, he was urged to think of some more responses. This was done to counteract the tendency of some Ss to "give up" readily. If he then gave any more responses, the above procedure was followed.

selected on the basis of diversity and overall representativeness of their parent batteries. As noted above, all were presented in a different randomized order. All procedures involve the generation of associations to specified stimulus requirements.

Torrance procedures (5). The four Verbal and one Figural (nonverbal) tasks included here are taken from the experimental edition of the Torrance Tests of Creative Thinking, Form A (1966). From several independent studies (e.g., Eherts, 1961; Wodtke, 1964; and Mackler, 1962), the technical manual summarizes test-retest reliability coefficients based on batteries including these tasks. Most of these coefficients, based on battery totals and on separate fluency, flexibility, and originality totals, are in the +.60's and +.70's , with a range from +.34 to +.97. Similar correlations are reported between Forms A and B.

1. "Ask and Guess--Asking" (AGA). In this task, the stimulus is a simple line drawing of an elf-like figure bending down on hands and knees apparently looking at his reflection in the water. The drawing is handed to S, and he is requested to generate questions which would give him more information about the pictured scene. Questions which can be answered by looking at the picture are not scored, although they are recorded without comment.

2. "Ask and Guess--Guessing Causes" (AGG). This task utilizes the same drawing, and for continuity, always followed "Ask and Guess--Asking." This time the S is required to generate possible causes or antecedents of the pictured scene. (A third in this series, "Ask and Guess--Guessing Consequences," was not

used, as it was similar in principle to the "Just Suppose Test," which was included.)

3. "Unusual Uses" (UU). This task is identical in nature to Wallach's "Alternate Uses," requiring the S to generate possible uses for a verbally specified object. A single item--cardboard boxes, is used.

4. "Just Suppose" (JS). The stimulus in this task is a verbally specified "improbable situation": "Just suppose that clouds had strings attached to them which hung down to earth." The S's imagination is assisted with a simple line drawing. His task is to generate possible consequences of this improbable situation.

5. "Repeated Figures" (RF). In this, the only nonverbal task, S is given thirty sets of parallel lines, and with his choice of pencil, crayons, or both, must generate pictures or objects, incorporating a pair of lines into each drawing. He is also asked to provide a title for each. In order to provide a fluency score, a time limit of ten minutes is imposed. In the NEV condition, timing was done inconspicuously with a wristwatch, and S was not aware he was being timed.

Wallach-Kogan procedures (3). Wallach & Kogan present both split-half and item-sum reliability data for their procedures. For the three tasks described here, split-half reliability coefficients range from +.51 to +.93, with an average of +.80, while average item-sum coefficients fall in the +.70's for each task.

1. "Instances" (I). In this task, S must give possible instances of a verbally specified class concept. The first two of the four class concepts presented by Wallach & Kogan in this procedure were used: "Name all the round things you can think of"; and "Name all the things you can think of that will make a noise." Examples of acceptable answers for "round things" are "ball," "globe," or "tree trunk"; while acceptable "noisy things" include "cars driving," "water running," or "hitting a baseball."

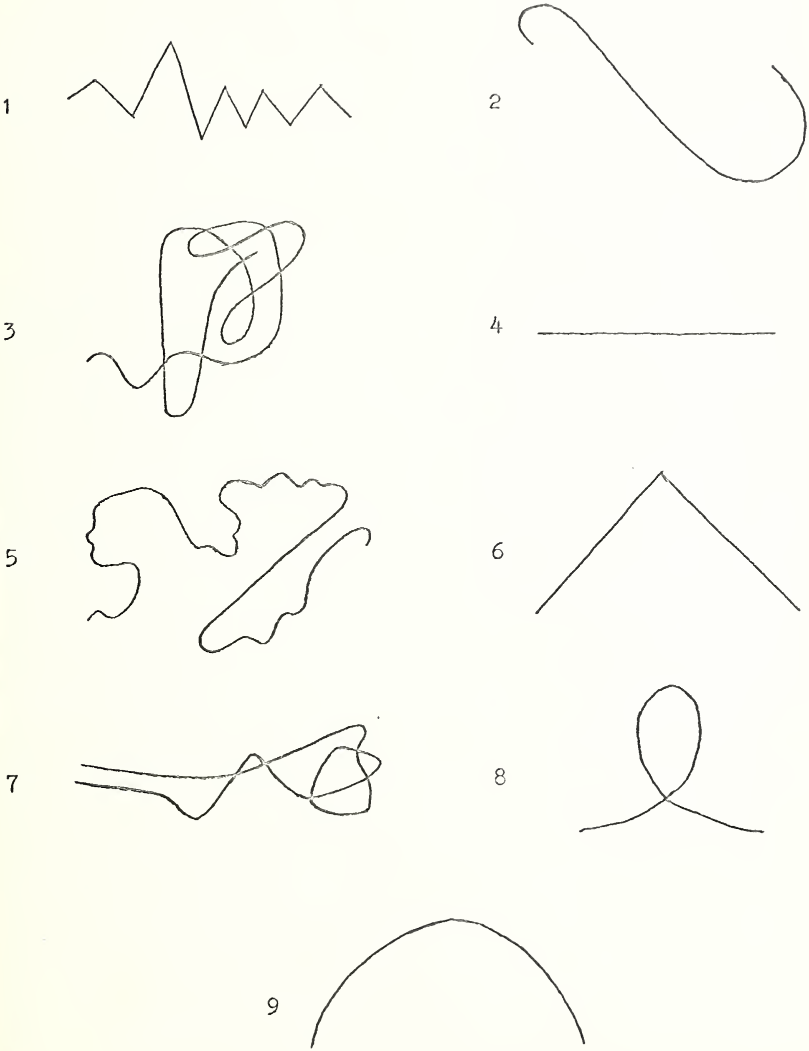
2. "Alternate Uses" (AU). This task requires the S to generate possible uses for a verbally-presented object. Of the eight items presented by Wallach & Kogan, four were selected at random: (1) an automobile tire--either the tube or the outer part; (2) a cork, or a piece of cork; (3) a button--the kind used on clothing; (4) a key--the kind used in doors.

3. "Line Meanings" (LM). This task utilizes nine visual stimuli in the form of abstract lines, and requires S to provide meanings or interpretations pertaining to the form of each line. Each line is a single unit, drawn with a black felt pen on a separate white 4 in. x 6 in. card. These lines are reproduced in Figure 2, numbered in the order of administration. All cards were sealed in clear plastic, so they could be handled easily. Turning the cards in various directions was freely encouraged.

Scoring

Specific scoring criteria, techniques, and rationales

Figure 2. "Line Meanings" Stimuli



are spelled out separately for Torrance and Wallach-Kogan tasks, as minor differences inevitably arose.

Torrance procedures. Torrance procedures were scored professionally by Torrance-trained scorers using acceptance criteria, flexibility categories, and originality norms published in Torrance (1966). Published inter-scorer reliability coefficients for all Torrance tasks generally exceed +.90.

Fluency was defined for each task as the number of acceptable responses produced. A response was accepted as appropriate if it conceivably met task requirements as set forth in the instructions. Where there was doubt, scoring appeared to be fairly strict. Approximately five percent of the responses to Torrance tasks in this sample were judged inadmissible by the scorers. Inadmissible responses included repetitions, those judged bizarre or irrelevant, and those which conveyed to the scorers that the S had failed to respond primarily to the task instructions.

The flexibility score for each task except "Just Suppose" was the number of different conceptual categories to which the given responses belonged. For example, if an S gave specific unusual uses for a cardboard box, such as "make a doghouse," "put magazines in," "put popcorn in," and "make a kite," he would receive a separate flexibility credit for each of the categories: "animal shelter," "container," and "toy." The flexibility score for "Just Suppose" was defined as the number of shifts from one category to another. Flexibility categories were evidently selected to be logical and to empirically fit about 98 percent of

available responses. No further criteria for category selection, level of abstraction, or breadth are provided in the Torrance manuals.

The originality score for each task was composed of a weighted rating for each acceptable response. Weights were based on statistical infrequency in available normative samples, and on the judged amount of mental effort which the response demonstrated. Originality norms for "Repeated Figures" were derived from a sample of 381 normal Ss from kindergarten through high school. Each response was assigned a weight of from zero to three. Responses occurring over twenty percent of the time in this sample, or which showed little "creative strength,"* were scored "0." A score of "1" was given to those responses occurring between five and nineteen percent of the time, and "2" to those occurring between two and four percent of the time. Responses occurring less than two percent of the time in the standardization sample, or which showed unusual "creative strength" in the scorer's judgment received a score of "3." These weights were summed for each response to yield the originality score for a given task. No standardization sample characteristics or cutoff points for originality weights were provided for the Verbal procedures; however, the identical principles were evidently followed.

The elaboration score, calculated in the present study

---

* "Creative strength" is defined by Torrance as the quality of a response which makes it "beyond what is learned, practiced, habitual, and away from the obvious and commonplace." (Torrance, 1966, Directions Manual and Scoring Guide, Verbal Test, Booklet A, Research Edition, p. 13.)

for "Repeated Figures" only, was defined as the number of ideas communicated by each picture, beyond the basic idea. An elaboration point was added for such extra details as color (when it added an idea), shading, decoration, or any extra meaningful line. This score was included to observe its relation to creativity and intelligence under varying task conditions.

Wallach-Kogan procedures. All scoring of the Wallach-Kogan procedures was done by the examiner. To control for possible scoring bias, E remained "blind" with respect to task condition. That is, raw data sheets were not identified with the assessment condition under which S was tested.

Fluency was again defined for each task as the number of different responses meeting task requirements. Since the Wallach-Kogan tasks consisted of two or more subtasks, the fluency score for each task represented the sum of acceptable responses to all subtasks. For example, the fluency score for "Line Meanings" was the sum of all acceptable responses to each of the nine line stimuli. Flexibility and originality scores were similarly derived. Inadmissible responses--repetitions or responses judged bizarre, inappropriate, or inconceivable in terms of task instructions--were extremely rare, occurring about one percent of the time.

The flexibility score for the Wallach-Kogan data was designed to follow that of Torrance as closely as possible. Since Wallach & Kogan did not score their data for flexibility, it was necessary to construct conceptual categories for all responses. A separate set of categories had to be devised for each subtask.

The guiding rationale was to follow the natural groupings of the responses where possible, so as to empirically and logically construct categories which were most meaningful, and which enabled most of the responses to be clearly classified with a minimum of overlap. For example, for the "Instances 1" task (round things), categories were chosen primarily according to the functions of round objects: e.g., art, educational aids, household items, transportation, etc. The categories for "Instances 2" (things that make noise), however, were based on the sources of particular sound qualities: e.g., animals, moving or sliding objects, sudden sharp impact, rotation on an axis, etc. Where a response could fit into more than one category, an arbitrary assignment was made and followed consistently. The few responses in each subtask which could not be classified were placed in a "miscellaneous" category.

A middle course was chosen in determining level of abstractness and breadth of categories. Too concrete and limited category scope would result in little distinction between fluency and flexibility, since the number of categories would approach the number of responses. On the other hand, too broad and abstract categories would fail to provide meaningful distinctions between Ss. In general, twenty to thirty categories per subtask appeared optimal; this number was consistent with Torrance's work. Undoubtedly many additions and substitutions of categories could have been made to fit the data equally well. In a few cases, arbitrary decisions created or eliminated questionable categories. It is felt, however, that changing the categories

slightly would not significantly alter the flexibility score, and certainly would not systematically favor one treatment group over another, since categories were formed with no knowledge of which responses were given by EV and NEV Ss.

No established originality norms were available for the Wallach-Kogan data; hence, it was decided to follow the authors' procedure of defining an original response as one which occurred only once in the experimental sample. With an N of 80, this represents occurrance 1.25 percent of the time in the parent population. The originality score for each task was the simple sum of all the original responses for its component subtasks. Torrance's concept of "creative strength" was not used, as it was desired to avoid reliance on subjective judgments wherever possible.

It was necessary in calculating originality to distinguish those responses occurring only once from those occurring more than once. Each different response was therefore placed on a separate 3 in. x 5 in. card, along with an identification of the subtask from which it came. Responses occurring more than once in the sample were indicated by a mark on the appropriate response card. Lists of original responses for each subtask were compiled from the remaining cards, and each S's share of the original responses represented his originality score.

The greatest single scoring problem encountered was the frequent necessity to judge whether two or more similarly worded responses to a single item in fact represented the same or different basic ideas. This judgment was crucial to originality

scoring, for two responses judged to be the same would be equivalent to two occurrences of a single response. Hence, neither response could be considered original. Occasional additional judgments had to be made concerning appropriateness of a response. These problems were handled by a scoring reliability check utilizing three independent judges. All responses for which the examiner could not make a definite judgment were taken from a randomly selected group of four out of the fifteen Wallach-Kogan subtests. Judges rated each response "appropriate" or "inappropriate" to task requirements, and each pair of responses in question as "same" or "different," based on a general scoring guide provided by the examiner. (See Appendix C.) Responses judged inappropriate by two out of three judges were excluded. Five responses were thus dropped. On the "sameness-difference" judgments, two out of the three judges agreed independently with the examiner's judgment on 75 percent of all pairs in question.* This was considered satisfactory evidence of inter-scorer reliability, as responses for which serious scoring questions were raised represented only about ten to fifteen percent of the total number of responses.

Principal Analyses

Dependent variables. The analyses described in the

---

* A number of "filler items"--responses selected and paired at random--were also inserted as a control. As predicted, these items elicited one hundred percent interjudge agreement.

"Results" section involved twenty-five basic creativity scores: fluency, originality, and flexibility scores for each of the three Wallach-Kogan and five Torrance procedures, plus an additional elaboration score for "Repeated Figures." The four RF variables, however, were omitted from the major correlational analyses described below, and were considered separately, because of the unique task requirements posed by this single figural procedure.* Hence, generalizations from these analyses are limited to the nine Wallach-Kogan and twelve Torrance Verbal scores. The remaining dependent variables were three IQ scores: verbal, nonverbal, and total IQ.

Analytical procedures. Results were considered separately for each hypothesis. In order to make correlational findings more comparable to previous research, it was decided to use two parallel methods of analyzing correlational patterns. Both are applications of the "common variance" approach. The first, used by Wallach & Kogan in their own study, may be termed the "mean r" method. It involves computing average r's by taking the simple mean of the correlations between all pairs of relevant variables. For example, the average correlation between Wallach-Kogan measures and IQ was found by taking the mean of the r's between each of the nine creativity scores and each of the three IQ scores.

The second, the "index and totals" method, combines the "common variance" approach with the traditional Torrance scoring

---

* RF is the only timed task, and the only one where S must draw without verbalizing his responses.

procedure. Overall creativity indexes were computed separately for Torrance and Wallach-Kogan procedures, by summing the twelve Torrance and nine Wallach-Kogan variables. Within each index, furthermore, total fluency,* originality, and flexibility scores were computed by summing across procedures. For example, the originality scores for I, AU, and LM procedures were summed to yield a total originality score within the Wallach-Kogan index. Correlations were then calculated involving relevant totals and index scores analogous to those obtained by the "mean r" method. The "index and totals" method enables the results of the present investigation to be compared with the findings discussed in the "Introduction." (pp. 12-14, 19-22)

The correlational analyses described here were instrumental in testing hypotheses 1a, 1b, 2a, and 3. Specifically, hypothesis 1a required the comparison of five coefficients comprised of fluency,* originality, and flexibility scores obtained under EV conditions. These five r's measured the variance shared (1) among Wallach-Kogan variables; (2) between Wallach-Kogan variables and IQ; (3) among Torrance variables; (4) between Torrance variables and IQ; and (5) between Wallach-Kogan and Torrance variables. Hypothesis 1b involved the same comparisons among variables obtained under NEV conditions. To test hypothesis 2a,

---

* Fluency scores used in computing correlations are replaced in the tables by "number" scores, derived by subtracting originality from fluency. The number score is a corrected fluency score, designed to eliminate the intrinsic correlation between fluency and originality scores, thus providing uncontaminated correlations between these two variables.

comparisons were made between the corresponding EV and NEV correlational patterns based upon flexibility scores alone vs. fluency and originality scores combined. For hypothesis 3, relevant creativity-intelligence correlations involving the Torrance variables were compared with those involving the Wallach-Kogan variables, holding constant the assessment condition and scoring method.

The remaining hypotheses were tested by standard analyses of variance. For hypotheses 1c and 2b, one-way analyses of variance comparing EV with NEV conditions were performed on all twenty-five creativity variables, plus subtotals. To test hypothesis 2b, resulting F values for flexibility scores were compared with those for fluency and originality scores. Two-way analyses of variance comparing middle with lower social class, and EV with NEV conditions on all twenty-five creativity variables provided the results for hypotheses 4a and 4b. Finally, a similar set of two-way analyses was performed on all creativity variables, comparing EV and NEV conditions, and high with low IQ groups. These groups were formed by dividing the total sample at the median total IQ score.

CHAPTER III.

RESULTS

Subject Characteristics by Assessment Condition

In order for meaningful comparisons to be made between creativity scores obtained under EV and NEV assessment conditions, it must be shown that Ss assigned to these two conditions did not differ systematically on any of the major selection variables. Table 2* presents the means and standard deviations of EV and NEV groups for each of the five selection variables: age, social position rating, verbal IQ, nonverbal IQ, and total IQ. One-way analyses of variance were performed to test the differences between means for significance. The data show that on none of the five selection variables did the two treatment groups differ significantly. It would appear, then, that any differences in creativity patterns which emerge between the two groups are not attributable to differences in age, social position rating, or intelligence.

---

* Appendix D contains all tables discussed in the present chapter.

Effects of Assessment Context

> Hypothesis 1a. Based on an average of fluency (number), originality, and flexibility scores, both the Torrance and Wallach-Kogan procedures given under evaluative (EV) conditions will correlate as much or more with IQ as with each other.

Table 3 presents the intercorrelations among the nine Wallach-Kogan creativity variables and between these variables and verbal, nonverbal, and total IQ (using the "mean r" method) for the forty Ss in the EV condition. It is evident that under EV assessment conditions, the Wallach-Kogan measures show considerable internal cohesiveness. Correlations range from +.07 to +.95, with an average r among the nine variables of +.52, which is highly significant. Of the 36 coefficients, 25, or 70 percent, are significant beyond $p=.01$, while two more are significant at $p=.05$.

Correlations between the Wallach-Kogan measures and intelligence, on the other hand, are uniformly low, ranging from -.04 to +.38, with a mean of +.10. Only two of the 27 r's reach significance at the .05 level. These results support the conclusion that, contrary to expectations, the IQ-independent creativity dimension found by Wallach & Kogan is not dependent on the presence of NEV assessment conditions, nor on the absence of a flexibility score.

The corresponding correlational patterns among the twelve Torrance Verbal creativity variables, and between these variables and verbal, nonverbal, and total IQ, are presented in Table 4.

The findings are quite different from those of Table 3. Correlations among the Torrance variables range from -.29 to +.76, with a mean of +.28. Analysis of the 64 coefficients reveals that 27, or less than half, reach the .05 level of significance, while 22 of these are also significant at the .01 level. Hence, *under EV conditions, the Torrance measures on the average do not appear to hold a significant amount of their variance in common.*

How do the Torrance measures relate to intelligence under EV conditions? Table 4 shows a range of correlations from -.20 to +.36, with an average of +.16. Of the 36 r's, only one is significant at p=.01, with two more significant at p=.05. Thus, *using the "mean r" method, the Torrance measures are about as unrelated to intelligence as are the Wallach-Kogan measures. Neither threat of evaluation nor presence of the flexibility score produced significant average intercorrelations between intelligence and either group of creativity variables.*

It has been shown that under EV conditions, the Wallach-Kogan measures describe a reasonably coherent dimension, whereas the Torrance measures do not. How, then, will the two sets of measures intercorrelate with each other? Table 5 presents these results. Averaging the 108 r's among all pairs of creativity measures yields a figure of +.29, which falls just short of significance at the .05 level. The range of r's is from -.09 to +.78. Forty-six, or 43 percent of the coefficients, reach significance at p=.05, while 28 of these are also significant at the .01 level. These figures are very similar to those for the Torrance measures alone (see Table 4), suggesting that *the Wallach-Kogan variables*

measure relatively little in common with the Torrance Verbal measures, under EV assessment conditions.

In short, when the "mean r" method is used to compare average correlations among the creativity and intelligence variables in question given under EV conditions, only the Wallach-Kogan measures intercorrelate significantly among themselves. The Torrance measures bear no significant relationship to one another, to the Wallach-Kogan measures, or to intelligence. However, their variance shared with the other creativity measures (r=+.28 and +.29) is somewhat higher than that shared with IQ (r=+.16). *Hypothesis 1a is not supported*, for all creativity variables show, on the average, less correlation with intelligence than with each other.

What is the effect on these correlational patterns if the "index and totals" method is substituted for the "mean r" method? Table 6 summarizes these findings. Considered in this way, all the intercorrelations among creativity variables are highly significant, while those between the creativity indexes and intelligence are not. *Whereas formerly only the Wallach-Kogan battery formed an internally coherent, intelligence-free dimension, this time the Torrance Verbal battery does so as well, although less strongly. Furthermore, the Wallach-Kogan and Torrance indexes are also strongly intercorrelated, suggesting that they may in fact tap the same basic dimension.* Evidently the pooling of separate creativity variables reduces the variance sufficiently to allow fairly stable dimensionality to emerge. It should be noted that *these results also fail to follow the*

pattern predicted in hypothesis 1a. The proposed creativity dimension emerges despite the presence of EV conditions and the flexibility component.

> *Hypothesis 1b.* Based on an average of fluency (number), originality, and flexibility scores, both the Torrance and Wallach-Kogan procedures given under nonevaluative (NEV) conditions will correlate as much or more among themselves and between each other as with IQ.

Using the "mean r" method, Table 7 presents the parallel NEV correlations among the nine Wallach-Kogan measures and between these measures and verbal, nonverbal, and total IQ for the forty Ss in the NEV group. It is clear that *the emergent correlational patterns are almost identical to those elicited by EV conditions.* The mean of the 36 creativity r's is +.59, with a range from +.36 to +.86. Every one of the coefficients is significant at p=.05, and all but two surpass the .01 level as well. Thus, *while the coherence of the Wallach-Kogan creativity dimension is manifest under either assessment condition, it appears to be slightly stronger under NEV conditions.*

Likewise, the Wallach-Kogan measures are similarly unrelated to intelligence under NEV conditions; they average +.11, ranging from -.17 to +.41. Three out of the 27 r's, or about 11 percent, attained significance at p=.05, with only one of these also significant at p=.01. It is evident from these data that *both the high intercorrelations of the Wallach-Kogan measures and their independence from convergent thinking measures are*

relatively unaffected by wide variations in objective assessment context as employed in this study.

Are the patterns among Torrance variables similarly unaffected by assessment conditions? Data in Table 8 suggest that this is generally the case. The average among the 64 creativity intercorrelations is +.28--the same figure obtained under EV conditions. The range is slightly less wide, from -.07 to +.77. Twenty-three r's, or 36 percent, are significant at the .05 level; eighteen of these reach significance at the .01 level. The comparable figures for the Torrance *vs.* IQ correlations show an average r of only -.03, with a range from -.39 to +.17. Only one of the 36 r's is significant at p=.05. Therefore, *NEV assessment conditions fail to elicit an increase in common variance among the Torrance measures, but do provide slightly greater independence from IQ.*

The final comparison of average correlations under NEV conditions is between the Wallach-Kogan and Torrance variables. These coefficients are presented in Table 9. Averaging the 108 r's yields a figure of +.27, as compared with +.29 under EV conditions. The range of correlations is from -.05 to +.68. About 40 percent (42) are significant at p=.05, with half of these also significant at p=.01. *As under EV conditions, then, the individual Torrance measures appear to relate to the Wallach-Kogan measures about the same as to each other; this relationship falls just short of significance at the .05 level.*

To summarize the findings based on the "mean r" method, the overall relationships between the nine Wallach-Kogan varia-

bles, the twelve Torrance variables, and intelligence, are presented in Table 10 for EV and Table 11 for NEV conditions. Table 10 summarizes Tables 3, 4, and 5, while Table 11 summarizes Tables 7, 8, and 9. It is clear that *under both assessment conditions, correlations between creativity and IQ measures are sufficiently low to allow the conclusion that neither Torrance nor Wallach-Kogan procedures were measuring intelligence*. Combining weighted average correlations for all 21 creativity variables with intelligence yields r's of +.13 and +.03, respectively. With respect to creativity intercorrelations alone, *the Wallach-Kogan measures show high and significant intercorrelations (+.52 and +.59), whereas the Torrance measures do not (+.28 for both EV and NEV)*. The similar failure of the two sets of measures to intercorrelate significantly is probably largely attributable to the lack of coherence among the Torrance measures, as shown in Tables 10 and 11.

Using the "mean r" approach on average correlations, then, *hypothesis 1a is not supported*, for even the Torrance measures are more highly intercorrelated than either creativity program is with IQ, under EV conditions. *Hypothesis 1b is supported*, however, for the same overall patterns prevail under NEV conditions. Based on an average of fluency (number), originality, and flexibility scores, therefore, *there is no evidence that the different assessment conditions had any significant effect on creativity-intelligence correlational patterns*.

Do the data based on the "index and totals" method point to the same conclusion? Table 12 presents these data, for the

NEV condition. *As under EV conditions, intercorrelations among creativity variables are highly significant, while creativity and IQ are minimally related in both sets of measures.* In comparison with the same figures for EV (Table 6), the r's among creativity totals are uniformly lower. At the same time, however, the r between the Torrance index and total IQ drops from +.30 to -.06 under NEV conditions. This finding supports the prediction that the Torrance measures are less apt to be contaminated by intellectual functions when administered under NEV conditions. On the whole, however, the correlational patterns are similar to those elicited under EV conditions; hence, *hypothesis 1b is supported*, as it is when one considers the data obtained by the "mean r" approach.

"*Repeated Figures.*" It was mentioned earlier that "Repeated Figures" was omitted from the preceding analyses because it imposed time restrictions and drawing requirements which were not present on the other creativity procedures. Do these additional requirements affect the relation of this task to the other creativity procedures or to intelligence? If they are of no consequence, one would expect the average correlation between the four RF variables and twelve Torrance variables to be similar to that among these remaining Torrance variables themselves; that between the RF and Wallach-Kogan variables to be similar to the mean r between Torrance Verbal and Wallach-Kogan variables; and that between RF variables and IQ to be similar to the mean r between Torrance Verbal and IQ measures. Table 13 presents the intercorrelations between RF variables and all other creativity

and IQ measures, for both EV and NEV conditions.

Under EV conditions, the average r between RF variables and the remaining Torrance variables is +.01, as compared with +.28 among Torrance Verbal variables alone (Table 10). For RF vs. Wallach-Kogan variables, the mean r is also +.01, compared with +.29 for Torrance Verbal vs. Wallach-Kogan measures. Finally, the mean r for RF variables vs. IQ is +.22, compared with +.16 between Torrance Verbal measures and IQ. The evidence suggests, therefore, that under EV conditions, RF does not measure the same thing that is measured by the remaining creativity procedures. On the other hand, its relatively strong relation with intelligence, while not significant, suggests that the effect of time restrictions and drawing requirements may be to tap intellectual, rather than creative abilities.

Is this pattern altered by NEV conditions where S was timed, but without his prior knowledge? From Table 13, the mean r between RF variables and Torrance Verbal measures under NEV conditions is +.06, compared with +.28 among Torrance Verbal variables alone (Table 11). For RF vs. Wallach-Kogan variables, and Torrance Verbal vs. Wallach-Kogan, the r's are +.21 and +.27, respectively. Finally, the RF measures intercorrelate with IQ on the average of -.07, compared with an average of -.03 for Torrance Verbal vs. IQ. Hence, under NEV conditions, RF behaves much more like the remaining Torrance measures in that it correlates more highly with the Wallach-Kogan measures than with IQ. Evidently, the unique requirements imposed by the RF task are of little consequence when the task is administered under NEV condi-

tions, but cause the RF variables to be somewhat more related to IQ under EV conditions. In neither case, however, does RF appear to tap abilities in common with the Torrance Verbal tasks. Hence, there appears little justification for pooling it in an index composed of Torrance Verbal procedures.

> Hypothesis 1c. All mean creativity scores will be higher under NEV than under EV conditions, across IQ and SES.

This hypothesis was tested by a one-way analysis of variance between EV and NEV conditions for each of the twenty-five creativity variables. Additionally, analyses of variance were performed on summed fluency, originality, and flexibility scores within Wallach-Kogan, Torrance Verbal, and Torrance Total (Torrance Verbal plus "Repeated Figures") tasks. The results are summarized in Table 14. Of the 34 comparisons, only one (RF-fluency) reaches significance at $p=.05$; this is in opposition to the predicted direction. Two other RF variables, flexibility and originality, approach significance ($p=.10$), also in opposition to prediction. Three other comparisons approach significance in the predicted direction--I-flexibility, JS-originality, and Wallach-Kogan total flexibility. However, these are hardly more than would be expected by chance. A closer look at the means suggests that, with the exception of the three RF means already mentioned, 29 of the 34 means are higher in the NEV condition. By chance, on the other hand, one would expect 17, or

half, to be higher in the NEV condition. A sign test on these 34 comparisons yields a chi-square value of 16.94, which is significant beyond the .01 level. A similar sign test including only the 25 basic creativity variables yields a chi-square value of 7.53, also significant at p=.01. Apparently, then, NEV conditions consistently facilitate creativity scores on the whole, although the facilitory effect is insufficient to attain significance. Hypothesis 1c is thus partially supported. Concerning the three RF variables which were apparently facilitated by EV conditions, it may be said at this point that such a finding is consistent with the earlier evidence (p. 59) that under EV conditions, RF is a more effective measure of IQ than of creativity. A possible explanation is considered in the "Discussion" section.

Effects of Scoring Method

> Hypothesis 2a. The correlational patterns predicted in hypotheses 1a and 1b will emerge strongly for fluency and originality together, but will fail to emerge for flexibility alone.

The aim of this hypothesis is to evaluate the effect of the flexibility score on creativity-intelligence correlational patterns. If the flexibility component of Torrance's scoring system is partially responsible for the past failure of his measures to form an IQ-independent creativity dimension, then the deletion of this measure from either the "mean r" or "index and totals" analyses should result in higher intercorrelations among remaining creativity variables.

For the EV group, Tables 15 and 16 summarize the overall mean creativity-IQ correlations for fluency (number) and originality variables together, and for flexibility variables, respectively. The correlations upon which these means are based are those presented in Tables 3, 4, and 5. Tables 15 and 16 thus represent a breakdown of Table 10. Inspection reveals that in no case does the average r for flexibility differ from the corresponding r for fluency and originality by more than +.05. Over all tasks, then, the deletion of flexibility changes the magnitudes and patterns of correlations very little, under EV conditions. Flexibility is neither more related to IQ, nor less related to the "creativity dimension" than are fluency and originality, under EV conditions.

Tables 17 and 18 represent the identical breakdown of Table 11 for the NEV group. Average correlations are based on those in Tables 7, 8, and 9. For NEV conditions, it is evident that there is almost as little differentiation between flexibility and the remaining creativity variables. In three cases (Wallach-Kogan measures, Wallach-Kogan *vs.* Torrance measures, and Torrance measures *vs.* IQ), the mean r for flexibility differs from that for fluency and originality by no more than +.06. The differences of +.12 for Torrance measures, and +.18 for Wallach-Kogan *vs.* IQ, suggest that flexibility tends to correlate slightly higher with both creativity and IQ measures under NEV conditions than do fluency and originality.

What is the role of flexibility in creativity-IQ correlations when the "index and totals" approach is used? This anal-

ysis is parallel to that recorded in Tables 6 and 12; however, the indexes used this time consist of only total fluency (number) and originality scores. Comparison of Table 19, which presents findings for both EV and NEV conditions, with Tables 6 and 12, reveals the effect of flexibility in the overall indexes. Comparing Table 6 with the EV correlations in Table 19, all corresponding r's are practically identical, except for the average r among Torrance totals. Interestingly, there is a higher, rather than a lower, intercorrelation when flexibility is included. When Table 12 is compared with the NEV correlations in Table 19, r's involving the creativity indexes are again nearly identical. Both Wallach-Kogan and Torrance totals, however, intercorrelate more highly when flexibility is included. In short, the exclusion of flexibility does not appear to affect the creativity indexes; however, the trend is for total flexibility to correlate more highly with total fluency and originality than these latter two do with each other. This finding, contrary to predictions, suggests that flexibility under either assessment condition is appropriate as a creativity measure, and does not increase the common variance between a total creativity index and IQ. The clear implication is that the flexibility component in Torrance's index is not an adequate explanation for the lack of significant variance shared in common by his measures.

> Hypothesis 2b. Mean fluency and originality scores will show a stronger increase from EV to NEV than will mean flexibility scores.

To test this hypothesis, we turn again to Table 14, which presents the analyses of variance across task conditions for mean fluency, originality, and flexibility scores for all procedures. The prediction would be supported if the F values for fluency and originality contained a greater percentage which attain significance than did the F values for flexibility. The only variable which attained even the .05 level of significance was RF-fluency, and this was opposite the predicted direction. _There is no evidence, then, that any measure--fluency, originality, or flexibility--is increased by NEV conditions significantly more than any other measure; therefore, hypothesis 2b is not supported._

Effects of Test Source

> _Hypothesis 3._ When similar assessment conditions and scoring methods are employed, no significant differences will emerge between Torrance and Wallach-Kogan procedures with respect to (a) intercorrelations among creativity variables, or (b) correlations with IQ.

The present investigation has undertaken to compare two assessment conditions (EV and NEV), and two alternate methods of scoring (fluency, originality, and flexibility; fluency and originality). Pairing each scoring method with each assessment condition yields four combinations of matched conditions under which Torrance and Wallach-Kogan procedures may be compared. These comparisons have already been made separately for each scoring method under each assessment condition. (See Tables 10, 11, 15-18).

Table 20 summarizes this information, comparing Torrance Verbal and Wallach-Kogan variables. The patterns within each combination of matched conditions are remarkably consistent. Regardless of assessment condition or scoring method, the Wallach-Kogan measures are the only ones to intercorrelate significantly. Under none of the four combinations do the Torrance Verbal measures intercorrelate significantly either within themselves or with the Wallach-Kogan measures; in all cases, however, significance approaches the .05 level. Looking at the creativity-intelligence correlations, in no case is there a significant relationship between Wallach-Kogan or Torrance and IQ measures. Nevertheless, the Torrance measures do appear to be somewhat more related to IQ under EV than NEV conditions. From these patterns it is clear that *the Wallach-Kogan measures possess an intrinsically high degree of common variance, which is independent of intelligence*, and which is virtually unaffected by the above manipulations of assessment context and scoring method. On the other hand, *the Torrance Verbal measures possess a much smaller degree of shared variance, which, although largely intelligence-free, is less so under EV conditions. The addition of the flexibility score has no significant effect upon any average creativity-intelligence correlations. Therefore, part (a) of this hypothesis must be rejected, while part (b) receives support*. One must consequently look beyond differences in assessment context and scoring method to account adequately for the failure of the Torrance variables to form an internally consistent creativity dimension, similar to the Wallach-Kogan dimension.

Effects of Socioeconomic Status

> Hypothesis 4a. Middle class Ss will obtain higher creativity scores than lower class Ss on verbal creativity procedures, regardless of assessment context.

Hypotheses 4a and 4b were tested with a two-way analysis of variance comparing social class and assessment condition for each of the twenty-five creativity variables. Middle class Ss (N=21) are those with a Social Position Rating of 1 or 2; lower class Ss (N=33) are those with a rating of 4 or 5. Table 21 presents the appropriate means and F values. Mean creativity scores for each social class are averaged over assessment condition. The F(SES) values in Table 21 represent the significance level of the difference between middle class and lower class mean creativity scores. Only one of the twenty-five comparisons, AU-originality, reaches significance beyond .05. On the whole, then, there are no significant differences between social classes on mean verbal or nonverbal (RF) creativity scores; hypothesis 4a is not supported.

> Hypothesis 4b. The effect of assessment condition on creativity scores will be the same for both middle class and lower class Ss.

This hypothesis is tested by the two-way analyses of variance presented in Table 21. The F(int) values in Table 21 represent the significance level of the interaction effect between

assessment condition and social class. The interaction effect measures the differential effect of assessment conditions on creativity scores across SES. A significant F(int) indicates that the two social classes do not respond in the same way to either or both assessment conditions. Table 21 shows, however, that none of the twenty-five F(int) values reach significance at any acceptable level. This finding confirms the hypothesis that both classes responded to the two assessment conditions in essentially the same way.

Effects of Intelligence

An additional set of two-way analyses of variance was performed to compare assessment condition and total IQ. High and low IQ groups were formed by dividing the total sample of eighty Ss at the median total IQ value. The mean IQ for the high group (N=39) is 117.0; that of the low IQ group (N=41) is 96.4. The purpose of these analyses is to determine (a) whether high IQ Ss differ from their low IQ counterparts on creativity scores in any consistent way; and (b) whether high IQ Ss react differently from low IQ Ss with respect to assessment condition. This approach is parallel to that described for SES, considering first the differences between means, and second, the interaction between intelligence and assessment condition. Table 22 presents the means and F values for high and low IQ groups on all twenty-five creativity variables. Inspection reveals no F(IQ) values which reach significance. This finding indicates that on the whole, Ss of higher

intelligence (mean IQ of 117) are neither more nor less creative than Ss of lower intelligence (mean IQ of 96.4). It corresponds to the earlier finding of minimal average correlations between creativity and IQ scores. On the other hand, four of the F(int) values are significant at the .05 or .01 levels, with four more approaching significance ($p = .10$). These values apply primarily to Torrance's "Asking" and "Guessing" tasks. A significant interaction indicates that the high IQ group achieves its highest creativity score under EV conditions, whereas the low IQ group does better under NEV conditions. It is to be expected that brighter Ss may not be aided as much by NEV conditions as are less bright Ss, for the former have already learned to be successful under EV conditions. In most cases, however, differences between high and low IQ groups are nonsignificant. This finding indicates that if creativity-intelligence correlational patterns had been considered separately for high IQ Ss (corresponding to the average IQ of Ss in most previous research), they would have differed little from the present sample as a whole.

To summarize, the results of hypothesis 1 suggest that differences in assessment context do not significantly affect creativity-intelligence intercorrelations. Furthermore, while NEV conditions do appear consistently to elicit somewhat higher creativity scores, these scores are not significantly higher. The findings from hypothesis 2 clearly indicate that flexibility is more appropriately considered a measure of creativity than of intelligence, regardless of assessment context. The results of hypothesis 3 suggest that regardless of assessment context

or scoring method, the Wallach-Kogan measures form a unified, IQ-independent creativity dimension. The Torrance measures, while generally unrelated to IQ, also fail to relate strongly to each other or to the Wallach-Kogan measures. The conclusions from hypothesis 4 indicate no significant differences in creativity scores between middle and lower class Ss, nor in their apparent perception of assessment conditions. Finally, it was also found that when Ss were divided into high and low IQ groups by a median split technique, no systematic differences existed between groups with respect to creativity scores or perception of assessment conditions.

CHAPTER IV.

DISCUSSION AND SUMMARY

Relation of Findings to Previous Research

This investigation began with the basic goal of reconciling the conflicting findings of the Torrance and Wallach-Kogan approaches to creativity assessment. The question was posed, "Why do the Wallach-Kogan measures form a dimension which is internally coherent and statistically independent from intelligence, whereas the Torrance measures do not?" Evidence was presented from several creativity studies which suggested that the practice of (a) task administration under IQ-test-like conditions, and (b) scoring for flexibility of conceptual categories (in addition to fluency and originality), were probably responsible for the failure of the Torrance measures to form the same intelligence-free dimension measured by Wallach & Kogan. The results of the present investigation, however, indicate that (a) the use of a gamelike, rather than a testlike, assessment context does not significantly increase either the IQ-independent dimensionality or the means of either the Torrance or Wallach-Kogan measures; and (b) on the whole, the flexibility score is not more intrinsically related to IQ or less related to creativity than are the

fluency and originality scores. They further indicate that the Torrance procedures (excluding "Repeated Figures"), correlate among themselves more than with IQ, but that the variance shared by individual Torrance measures nevertheless falls short of significance at the .05 level. This section aims to relate these results to the evidence cited in the "Introduction" section, and to provide alternate explanations for the discrepancy between the correlational patterns of the Torrance and Wallach-Kogan approaches.

Role of assessment context. Specifically, it was predicted that NEV conditions would produce results differing from EV in two ways. First, it was thought that all creativity measures, regardless of source, would intercorrelate more highly among themselves and less with IQ in the NEV context. Second, it was postulated that all creativity variables (or at least fluency and originality scores) would have higher means under NEV conditions. It was found, however, that the correlational patterns predicted for NEV occurred under both conditions. Where differences emerged between results for EV and NEV, the trend was usually in the predicted direction. For example, the correlation between Torrance measures and IQ was lower under NEV than EV conditions. At the same time, it was found that the vast majority of creativity means were higher under NEV, as predicted, but that practically none of these differences were significant at $p = .05$. The implication is that NEV conditions produced a consistent facilitative effect on creativity scores, but that this effect was too small to produce significant differences.

One explanation for these findings is that Ss failed to perceive the two assessment contexts as being as different as had been expected. Manipulation of "assessment context" actually involved manipulation of external cues--task instructions and the demeanor of the examiner. It was expected, as stated in assumption 3 (p. 25), that Ss would respond to these cues with different internal sets under EV and NEV instructions. Many uncontrollable intervening variables may, however, have "diluted" the differential response. For example, most Ss may have perceived the assessment situation as basically evaluative, regardless of instructions or E's demeanor. It is possible that by the sixth grade, most school children have developed generalized expectations of evaluation for activities within the school setting. Another possibility is that by the sixth grade, most children are maximally motivated to achieve in any new or ambiguous situation. In Allport's language (Allport, 1961, p. 229), achievement motivation has become "functionally autonomous." The child who has learned for six years that he is expected to perform at his best in the school setting may have difficulty altering this approach in response to a simple set of task instructions. The consistency of the Wallach-Kogan creativity dimension across assessment conditions, furthermore, allows the speculation that Wallach & Kogan's Ss may have approached their assessment situation as if it were evaluative.

Another factor which may have reduced the difference in results across assessment conditions is the absence of time limits under either condition. In the traditional Torrance adminis-

tration, rather strict time limits (five or ten minutes) were imposed on all procedures, apparently to facilitate group administration. In the present study it was desired to suggest time pressure by providing a stopwatch in S's view and by indicating that he would be timed. It was necessary, however, to equate the amount of time actually allowed under EV and NEV conditions; otherwise, differences in results could have been due to differences in time allotted the two groups, rather than to the presence or absence of evaluative pressures. Hence, it may be actual time restriction, rather then mere threat of time pressure, which has contributed to a lack of creativity dimensionality in previous Torrance research. This notion receives support in Mednick's (1962) formulation, discussed earlier, in which he postulates that the creative person's unique responses do not come forth until he has had time to exhaust his common ones. It could be tested by comparing timed vs. untimed groups under individual evaluative assessment conditions.

It was noted through inspection of the data that many creativity variables possessed rather high standard deviations relative to their means (See Table 14); this may represent an additional explanation for the absence of significant differences between EV and NEV conditions. The effect of a large standard deviation is to increase the overlap between comparison groups, thus lowering the F value and reducing the possibility of obtaining a significant difference between means. To compare the present standard deviations with those in previous research, Wallach & Kogan found S.D.'s (for boys) of 7.44 and 3.91 for "Instances"-

number and originality, compared to 20.42 and 7.71 (averaged across assessment condition) in the present study. Likewise, in the "Line Meanings" task, Wallach & Kogan's S.D.'s for number and originality were 13.91 and 6.08, compared to 16.80 and 19.79 in the present study. By contrast, however, the corresponding figures for "Alternate Uses" were lower in the present study (6.22 and 7.59) than in the Wallach-Kogan research (19.54 and 9.58). Similar data for the Torrance Verbal tasks are presented for total fluency, originality, and flexibility scores, based on a sixth grade sample. (Torrance, 1966, Norms-Technical Manual, Table 5.7, p.66) The standard deviations presented are 28.9 for fluency, 17.6 for originality, and 9.2 for flexibility. Corresponding figures for the present study are 8.58, 3.85, and 3.02, averaged across assessment condition. It is evident, then, that relatively large standard deviations may contribute to the lack of strong differentiation between assessment conditions only in the case of certain of the creativity variables, especially those in the "Instances" and "Line Meanings" tasks.

Role of flexibility. The finding that flexibility varied similarly to fluency and originality was likewise an unexpected one. Specifically, it was found that, in taking the average of all relevant intercorrelations, there were no differences under EV conditions. Under NEV conditions, both the Wallach-Kogan vs. IQ, and the Torrance Verbal intercorrelations were somewhat higher for flexibility than for the combination of fluency ("number") and originality scores. Using the "index and totals" method, which is the method previously used with Torrance measures, EV

again produced no differences; NEV elicited total flexibility scores which were more highly related to total fluency and originality scores than these latter two were to each other. Hence, if anything, flexibility enhances slightly the common variance among the creativity measures, and in no case relates more with IQ than with other creativity measures. Additionally, it was found that mean flexibility scores were similar to fluency and originality scores in appearing slightly, but not significantly, higher under NEV conditions. These findings support the conclusion that flexibility is not intrinsically related to IQ; moreover, this measure appropriately is considered as one of Guilford's divergent thinking factors. This implies that the "nexus of creativity" may include the ability to categorize as well as simply to generate appropriate associations.

This conclusion clearly is contrary to the findings reported earlier. Can this discrepancy be accounted for? As in the discussion on assessment context, the effects of timed, group administration on flexibility intercorrelations cannot be ascertained at present. There is no basis, however, for supposing that these effects would produce stronger correlations between flexibility and IQ than between IQ and the other creativity measures. Therefore this finding must at present remain unexplained.

Role of test source. It was predicted that under similar assessment conditions and with similar scoring methods, no differences would emerge in the creativity-intelligence correlational patterns of Torrance vs. Wallach-Kogan measures. It was found, on the other hand, that Torrance Verbal measures correlated among

themselves and with the Wallach-Kogan measures far less than the Wallach-Kogan measures correlated with each other, but somewhat more than these Torrance measures correlated with IQ. Since these patterns are relatively unaffected by assessment condition, exclusion of the flexibility score, or comparison by the "totals and index" approach, rather than the "mean r" approach, it is possible that the Torrance tasks measure a wider variety of abilities than do the Wallach-Kogan tasks. Furthermore, these abilities appear not to be highly intercorrelated, although they are relatively independent of intelligence, particularly under NEV conditions.

The question of whether these tasks all measure creative abilities is difficult to answer from the available evidence. According to our "common variance" notion, discussed earlier, these measures should share significant and substantial variance in order to be considered measures of a common dimension. Findings indicate that their shared variance based on average correlations of all pairs of variables falls just short of significance at the .05 level. Their shared variance based on summed total fluency, flexibility, and originality scores _is_ statistically significant, although still less than that of the Wallach-Kogan measures. One must ask, "What about the Torrance Verbal procedures prevents them from intercorrelating with each other and with the Wallach-Kogan tasks to the same degree that the Wallach-Kogan tasks intercorrelate among themselves?"

An important possibility is that the average intercorrelations are being lowered, not by a particular scoring category

such as flexibility, but rather by a particular task which measures something quite different from what is measured by the remaining tasks. This possibility was anticipated in the decision to remove "Repeated Figures" from the Torrance intercorrelations on the grounds that it posed unique task requirements. Inspection of Tables 4 and 8 suggests that the "Just Suppose" activity likewise bears minimum relationship to the remaining Torrance Verbal procedures. If JS is removed from Table 4 (Torrance Verbal intercorrelations under EV conditions), the average of the remaining 36 Torrance intercorrelations rises from +.28 to +.40. On the other hand, the average of the 27 r's between JS variables and remaining Torrance variables is only +.13. The parallel figures under NEV conditions (from Table 8, p. 55) are +.33 among the three remaining Torrance procedures, and +.20 for the average r between JS and the remaining Torrance procedures. Hence, under either assessment condition, the removal of JS results in a significant amount of shared variance among the remaining Torrance tasks. Is this variance also shared with the Wallach-Kogan tasks? Averaging the coefficients of the AGA, AGG, and UU procedures with those of the three Wallach-Kogan procedures yields figures of +.29 and +.28 for EV and NEV. These figures, which are the same as those between all four Torrance Verbal procedures and Wallach-Kogan procedures, indicate that even without JS, the Torrance measures only partially tap into the Wallach-Kogan dimension.

The evidence suggests, then, that the "Just Suppose" measures assess something other than what is measured by the remain-

ing Torrance procedures.* Further analysis indicates, however, that this "something other" is not intelligence. The average r between JS variables and IQ is +.05 for EV and +.07 for NEV. In short, then, it becomes difficult to know just what the Torrance tasks measure. Since one cannot rely on the construct validity evident in the formation of a unified dimension, one is left with some sort of face validity claim that the Torrance measures do indeed measure creative abilities.

An adequate explanation for *why* the Torrance procedures fail to show the degree of intercorrelation typical of the Wallach-Kogan tasks still has not been provided. A look at some subtle differences in scoring procedures may provide some clues. These differences occurred because of a desire to reproduce the scoring techniques of the respective authors as closely as possible.

One difference, mentioned earlier, involves the scoring of originality. Due to the lack of established norms, it was necessary to score the Wallach-Kogan tasks for originality by considering as original any response which occurred only once in the sample. Torrance tasks were scored for originality with a

---

* One would not expect, of course, to find extremely high r's among measures which tap a common dimension if the measures are to be useful in tapping different aspects of this dimension. For example, one would not expect to find an r of +.90 between Arithmetic and Vocabulary subtests on the WISC, for they obviously tap different functions. Nevertheless, they are moderately and significantly correlated, for they both tap the dimension of intelligence. Likewise, for JS to tap the creativity dimension, it should correlate moderately and significantly with other creativity measures.

weighted rating of 0 to 3, based on infrequency and "creative strength" in an unclearly defined normative sample. It was assumed that these scoring methods would be approximately equivalent, since both were based primarily on infrequency of occurrence. It is possible, however, that lack of equivalent weighting systems and normative samples reduced correlations between Torrance and Wallach-Kogan originality variables. To shed light on this possibility, the average correlations between Torrance Verbal and Wallach-Kogan originality variables were calculated from Tables 5 and 9. The average among the 12 r's for EV (Table 5) is +.24; for NEV (Table 9) it is +.23. These figures are only slightly lower than those between Torrance and Wallach-Kogan measures as a whole (+.29 and +.27), suggesting that these scoring differences are of little consequence. Likewise, it is possible that the Torrance corrected fluency score (fluency-originality) is less related to the other measures because of the use of originality weights which artificially inflate the originality score and hence reduce the corrected fluency score.

Another difference in scoring is that the acceptance requirements for a response appeared to be somewhat stricter in the Torrance than the Wallach-Kogan tasks. It was mentioned earlier that about five percent of the Torrance responses were discarded, as compared with about one percent of the Wallach-Kogan responses. This has the effect of reducing the number of acceptable Torrance responses given by an S relative to the number of Wallach-Kogan responses, and thus proportionately reducing the originality and flexibility scores as well. While this would

not affect the intercorrelations among Torrance or among Wallach-Kogan variables, it would decrease the average intercorrelations between the two sets of measures. Hence, the possibility that the Torrance tasks in fact do measure the same dimension that the Wallach-Kogan tasks do cannot be ruled out.

A final possible explanation concerns the problem of scoring reliability of the Torrance measures. Despite the high scoring reliability coefficients reported in the Torrance manual (Torrance, 1966), it appeared that no such reliability was achieved by professional scorers employed in this study. While it is difficult to prove that the scoring reliability for the Torrance measures in this study was actually lower than the published figures, lowered scoring reliability would likely be reflected in lower-than-expected intercorrelations among Torrance creativity variables.

"Repeated Figures"--its relation to other measures. It was mentioned that RF was excluded from the main correlational analyses because this task appeared to pose different requirements from the other procedures. The findings from Table 13 (pp. 59-60) indicate that under EV conditions, which are typical of Torrance administration, RF is related more highly to IQ than it is to either Torrance or Wallach-Kogan procedures. Its effect, then, is to increase the overall amount of variance held in common between Torrance measures and IQ, and to decrease the average variance held in common among the Torrance measures. Under NEV conditions, the inclusion of RF would not have affected the relationship of the Torrance measures with IQ, but would

have likewise decreased the common variance among the Torrance measures.

A look at Table 14 provides an explanation. It shows that RF variables were generally higher under EV than NEV conditions, while nearly all other creativity variables showed the opposite trend. RF was, moreover, the only procedure in which EV instructions told S to work as quickly and carefully as possible, because of a ten minute time limit. Apparently the inclusion of a specified time limit provided sufficient motivation that Ss as a whole did better on this task under EV conditions. The largest increase under EV was in the fluency score, suggesting that Ss drew more quickly so as to make more, rather than more elaborate, drawings. The higher relationship between RF variables and IQ under EV conditions lends further support to the notion mentioned earlier that the imposition of strict time limits tends to bring forth intellectual, rather than creative, abilities on so-called creativity tests. A further conclusion is that RF does not possess sufficient variance in common with the Torrance Verbal measures to be pooled legitimately into an index with the Verbal measures.

The relation of the "Elaboration" variable in RF to other variables is unclear in the present research. Elaboration is not related significantly to Wallach-Kogan or remaining Torrance variables (avg. $r = +.12$ and $+.15$ for EV and NEV, respectively). These figures are, however, somewhat higher than the average correlation of +.01 between all four RF variables and the remaining creativity variables, for both EV and NEV. A look at

the scoring inconsistencies for this measure suggests that this lack of relationship may be due to very low scoring reliability.

Role of socioeconomic status. A secondary purpose of this investigation was to compare middle and lower classes on mean creativity scores. Since all tasks but "Repeated Figures" required facility in verbal expression and association, it was predicted that middle class Ss would do better than lower class Ss. The results showed a trend in this direction, but no significant differences across classes were found. What does this finding mean? It may mean, for one thing, that within the sample tested, there were not sufficient class differences for strong differences in creative performance to be apparent. For example, no chronic welfare recipients were included in the sample. Furthermore, by the sixth grade, school may exert a "levelling effect," reducing the discrepancy in verbal and other skills between the social classes. On the other hand, it is possible that since the Ss were not required to write anything down, and since the tasks required a minimum of expressional skill, that class differences in expressional skill were irrelevant to this form of creativity assessment. One cannot, from the available evidence, determine whether the trend in favor of the middle class was due to differences in creative associational abilities, or to differences in expressional capabilities.

With regard to the nonsignificant interaction between social class and assessment context, one of two things may be indicated. On one hand, both social classes may have responded the same way to each assessment condition, as predicted by hy-

pothesis 4b. On the other hand, the difference between assessment conditions may have been insufficient to cause real differences in social class to emerge. In either case, there is no evidence that either class perceived assessment conditions differently.

Role of intelligence. The relationship of IQ to creativity variables was measured in two ways: by intercorrelation with creativity measures and indexes, and by a series of two-way analyses of variance on creativity variables comparing high *vs.* low IQ groups (divided at the median IQ for the total sample) and EV *vs.* NEV assessment conditions. As predicted, the Wallach-Kogan measures were related negligibly to IQ under either assessment condition. Contrary to previous findings, however, the Torrance measures also showed relatively low correlation with IQ. With the flexibility score included, four estimates of this relationship were obtained: average r's between variables and IQ; and the correlation between total index and IQ, for both EV and NEV. For EV, these figures are +.16 and +.30; for NEV, they are -.03 and -.06, as was shown in Tables 10, 6, 11, and 12. While all these r's are nonsignificant, the relation between Torrance Verbal variables and IQ was reduced under NEV, as predicted. The real question, however, is, "Why don't the Torrance Verbal measures intercorrelate significantly with IQ when scored for flexibility and given under EV conditions?" These conditions represent the way the Torrance procedures have most often been given in the past. This finding indicates that although the Torrance procedures were not clearly measuring creativity,

neither were they acting as inferior measures of intelligence, as earlier evidence had suggested (e.g., Yamamoto, 1965; Wodtke, 1964). Two possible reasons for this discrepancy are immediately apparent. One is that something about individual assessment conditions kept intellectual abilities from coming to the fore. The other, mentioned earlier as a possible factor in the lack of distinction between EV and NEV, is that strict time limits were absent, even under EV conditions. The effects of both group *vs.* individual administration and timed *vs.* untimed conditions upon creativity-intelligence patterns remains to be examined.

The results of the two-way analysis of variance between IQ and task condition suggest that there were no significant differences between high and low IQ groups on mean creativity scores. The implication is that even if the mean IQ in this sample had been considerably higher, results would have been very similar. The results further indicate an interaction effect on the Torrance "Asking" and "Guessing" tasks, such that the brighter Ss appeared to do better under EV conditions, while the duller Ss did better under NEV conditions. These results are in accord with findings by Wallach & Kogan that Ss who are high in creativity but low in intelligence seem to require a more supportive, evaluation-free task context than do Ss who are high in both creativity and intelligence. This finding furthermore provides a possible lead in the differential diagnosis of pseudoretardation *vs.* true mental retardation. If performing relatively better on certain creativity measures under EV conditions is a func-

tion of intellectual potential, rather than artificially lowered intellectual performance, one could hypothesize that pseudoretarded Ss would perform similarly to normal bright Ss; that is, they would do relatively better under EV than NEV conditions, while genuinely retarded Ss would be hampered severely by EV relative to NEV conditions. One would of course need to first determine the effect of assessment conditions on other variables associated with pseudoretardation, such as anxiety threshold, reasons for the pseudoretardation, etc. It should be remembered also that the Ss in this study all had an IQ of 80 or above, and hence none can be considered retarded in the usual sense.

Implications for Further Research

Methodology. A major aim of the present study was to refine existing research methodology to provide more meaningful assessment of creative abilities. What do the findings indicate regarding the methodological aspects of future research? It is apparent, first, that the Torrance Verbal tasks used in this study may be considered maximally independent of IQ when administered individually in untimed, game-like assessment conditions. In light of evidence cited in the "Introduction," that the Torrance tasks are correlated moderately with IQ when administered under timed, group testing conditions, researchers comparing Torrance creativity and IQ should consider the gains in interpretability of findings to be found in individual nonevaluative assessment. The next important methodological questions to be

answered with respect to both the Torrance and Wallach-Kogan tasks concern the possible effects of strictly timed _vs._ untimed, and group _vs._ individual assessment conditions. Does strict timing and/or group pressure account any better for the traditional lack of dimensionality of the Torrance measures than do the evaluative pressures used in the present study? Again, would the Wallach-Kogan dimensionality be decreased by strict timing or group administration? Studies designed to answer these questions could shed further light on optimum creativity assessment conditions.

Certain implications about the legitimacy of combining scores into an overall index are apparent. Since the Wallach-Kogan measures intercorrelate significantly, they may be combined legitimately into an estimate of creative functioning, analogous to IQ. The Torrance data indicate, on the other hand, that while AGA, AGG, and UU share sufficient common variance to be pooled into such an index, JS and RF do not. If RF is representative of the other figural tasks in the Torrance battery, then, there is no justification for combining Verbal and Figural scores to form a composite index. It might be found, however, that applying the "common variance" approach to each of the Torrance tasks, including those not used in the present study, would yield data allowing the formation of useful indexes _separately_ for Verbal and Figural procedures.

Another methodological possibility involves the application of a rotated factor analysis to the correlational matrixes obtained in the present study. Such an analysis, carried out

within the "common variance" framework, could shed further light on which Wallach-Kogan and Torrance procedures are the most efficient assessors of the "nexus of creativity," by showing which load highest on a hypothesized creativity factor. Additionally, the differences in emergent factor patterns between assessment conditions could reflect the effect of assessment condition on each specific creativity variable. Such an analysis would be more precise and sophisticated than the "average correlation" approach used in the present study.

Finally, this study suggests the need for a critical examination of the Torrance scoring procedures, with an effort toward making them more consistent and reliable. Despite high scoring reliability figures published by Torrance (1966), his scoring manuals provide insufficient information with which to score many responses. In particular, criteria for response acceptance and originality rating allow for considerable ambiguity and consequent guessing on many responses. This study raises the possibility, then, that the potential usefulness of the Torrance materials is hampered presently by a lack of sufficiently clear and detailed scoring criteria.

Populations. The present study involved only white boys in the public school sixth grade. Future research, however, is needed on several other populations. One of the biggest needs presently is for adequate creativity age norms for both sexes, analogous to those already developed for intelligence. For the Wallach-Kogan tasks, such an undertaking would make unnecessary the present tedious process of counting all responses occurring

only once in the sample to obtain the originality score. For the Torrance tasks, it would contribute greatly to the completeness and usefulness of present originality norms.

Another potentially fruitful population for creativity research is the "culturally deprived." If the present methods of creativity assessment are sufficiently free of dependence on expressional skills lacked by such children, they could prove very useful in identifying children who have latent thinking skills that might otherwise go unnoticed and undeveloped. It would also be interesting to see whether the dimensionality of the Wallach-Kogan measures holds up in a homogeneous "deprived" sample.

Finally, an exciting application of creativity assessment might be with emotionally disturbed children. Creativity measures could be used both as "projective" devices (by analyzing content), and as idiosyncratic indexes of changes in mental functioning as treatment progressed. The qualitative and quantitative relationship of creative abilities to emotional disturbance in children remains a fascinating, but relatively unexplored area, to which the measures in the present study may have great relevance.

Summary

This investigation developed out of a need to examine, and perhaps reconcile, conflicting findings arising from two major creativity assessment approaches with children. Specifi-

cally, previous research indicated that creativity instruments devised by Wallach & Kogan formed an internally coherent, intelligence-free dimension, whereas a similar set of instruments devised by Torrance did not. Evidence suggested two possible explanations for this discrepancy. First, the Wallach-Kogan procedures were administered under relaxed, game-like assessment conditions, while the Torrance procedures were given typically in a context similar to that of a school achievement test. Second, although both sets of tasks were scored for fluency (number of appropriate responses) and originality (number of statistically unique responses), the Torrance tasks received an additional score for flexibility (number of different conceptual categories used).

Based on past evidence, it was hypothesized that: (1) both sets of creativity measures would intercorrelate more highly with IQ than with each other's variables under evaluative (test-like) assessment conditions; (2) both sets would intercorrelate more highly with each other's measures than with IQ under non-evaluative (game-like) conditions; (3) flexibility would prove intrinsically more related to IQ and less related to the remaining creativity variables; and (4) under similar assessment conditions and scoring methods, no differences in creativity-IQ correlational patterns would emerge between Torrance and Wallach-Kogan measures. It also was predicted that nonevaluative conditions would elicit higher mean creativity scores than would evaluative conditions, and that middle class Ss would achieve higher creativity scores than would lower class Ss.

Forty white, sixth grade males, heterogeneous with respect to IQ and SES, were administered individually representative Torrance and Wallach-Kogan procedures under evaluative (EV) conditions, while an additional forty Ss received the same tasks under nonevaluative (NEV) conditions. Under EV conditions, task instructions placed a premium on correctness of responses, and a stopwatch in S's view provided implied time pressures. (Time under EV and NEV was actually kept equivalent.) Under NEV conditions, the procedures were presented as games; special efforts were made to achieve good rapport, time pressure was absent, and any hint of evaluation was displaced from S to the "games."

Results were analyzed principally by comparing averages of all pairs of correlations among relevant variables, and by standard analyses of variance. The major findings were:

1. Wallach-Kogan measures formed an IQ-independent creativity dimension regardless of assessment condition or presence of a flexibility score.

2. The Torrance Verbal procedures tended in the direction of a unified creativity dimension under either assessment condition and with or without flexibility, but this tendency did not quite reach significance when considering the average of r's between all pairs of variables.

3. The Torrance Verbal procedures shared the same amount of variance with the Wallach-Kogan procedures as they did with each other.

4. Neither threat of evaluation nor presence of the flexibility score produced significant average intercorrelations

between intelligence and either Torrance or Wallach-Kogan variables. The Torrance measures were somewhat less related to IQ under NEV than under EV conditions. In either case, however, they shared more variance with each other than with IQ.

5. The Torrance Figural task, "Repeated Figures," is not appropriately included with Torrance Verbal procedures in an overall index, because it was in general more highly related to IQ than to the Torrance Verbal creativity variables. Particularly under EV conditions, the effect of including a specific time limit was to draw forth intellectual-like abilities.

6. The effect of NEV conditions was to produce consistently higher creativity scores for nearly all procedures except "Repeated Figures," but the increases from EV means were, on the whole, nonsignificant.

7. The flexibility score was not intrinsically more related to IQ than were fluency or originality scores. Flexibility was, in fact, related to fluency and originality as much or more than these latter scores were to each other, especially under NEV conditions.

8. Flexibility did not differ from fluency and originality with respect to increases in means from EV to NEV conditions.

9. No significant social class differences emerged on creativity means or reaction to assessment conditions.

10. There was no difference between high and low IQ Ss with respect to mean creativity scores. However, on Torrance's "Asking" and "Guessing" tasks, high IQ Ss tended to do better in EV, while low IQ Ss tended to do better in NEV conditions.

APPENDIXES

Appendix A

HOLLINGSHEAD'S TWO-FACTOR INDEX OF SOCIAL POSITION (MODIFIED)

Occupation:

1. Higher executives, proprietors of large concerns (over $100,000), and major professionals, including military officers of Major or above, CPA's, architects, portrait artists, veterinarians.

2. Business managers, proprietors of medium-sized businesses ($35,000 to $100,000), and lesser professionals, including branch managers, office men, librarians, Lieutenants and Captains in military, nurses, social workers, public school teachers.

3. Administrative personnel, minor professionals, and small independent businesses ($6,000 to $35,000), including private secretaries, chief clerks, purchasing agents, sales representatives, chain store managers, Master Sergeants, interior decorators, lab assistants, morticians, surveyors, photographers, reporters.

4. Clerical and sales workers, technicians, and owners of little businesses (under $6,000), including railroad conductors, lab technicians, draftsmen.

5. Skilled manual employees, including barbers, foremen, policemen, auto mechanics, city firemen, railroad firemen, loom fixers, painters, toolworkers, farm owners under $10,000.

6. Machine operators, truck drivers, and semi-skilled employees.

7. Unskilled employees.

Education:

1. Professional degrees, including M.A., LL.B., B.D., and M.Ed.

2. Four-year college graduate

3. One to three years college; also business school

4. High school graduate

5. Ten to eleven years schooling

6. Seven to nine years schooling

7. Under seven years schooling

Weights:

Occupation - 7

Education - 4

Classes:	Social Position Scores:
I	11-17
II	18-31
III	32-47
IV	48-63
V	64-77

Appendix B

CREATIVITY TASK INSTRUCTIONS

EV Condition

Torrance procedures (5).

1. "Ask and Guess--Asking" (AGA). "This test and the next one will be based on this picture." (Hand picture to S.) "This one will give me a chance to see how good you are at asking questions to find out things that you don't know. Look at the picture. Try to figure out what is happening. What can you tell for sure, and what do you need to know in order to understand what is happening? Your job is to make up all the questions you can think of about this picture which you would need to ask to know for sure what is happening. Ask only questions which cannot be answered by just looking at the drawing, and think of the best questions you can. Tell me each question so I can write it down. Do you understand? All right, go ahead."

2. "Ask and Guess--Guessing Causes" (AGG). (This task always followed AGA.) "Now look at the picture again, and instead of asking questions, this time guess as many possible causes as you can for the action shown in the picture. In other words, think of the different things that might have led up to what is

happening in the picture. You may use things that might have happened just before the things happening in the picture, or something that happened a long time ago that made these things happen. Make as many guesses as you can--the more the better--and make the best guesses possible. Ready, begin."

3. "Unusual Uses" (UU). "Here's a(nother) test to see how well you can think of many different unusual uses for the same thing. Most people throw away their empty cardboard boxes, but these boxes have thousands of interesting and unusual uses. List as many of these interesting and unusual uses as you can think of. Do not limit yourself to any one size of box. You may use as many boxes as you like. Do not limit yourself to the uses you have seen or heard about; think of as many possible new uses as you can, and try to think of things nobody else would think of. Okay, begin."

4. "Just Suppose" (JS). "You will now be given an improbable situation--one that will probably never happen. You will have to just suppose that it has happened. This will give me a chance to test your imagination to see how well you can think out all of the other exciting things that would happen if this improbable situation were to come true. In your imagination, just suppose that clouds had strings attached to them which hung down to earth. Now, what are all the other things that would happen because of this? How would it make things different? Make as many guesses as you can."

5. "Repeated figures" (RF). "This test is to see how many objects or pictures you can make from the pairs of straight

lines below and on the next two pages. The pairs of straight lines should be the main part of whatever you make. With pencil or crayons, add marks to the pairs of lines to complete your picture. You can place marks between the lines, on the lines, and outside the lines--wherever you want to in order to make your picture. Try to think of things that no one else will think of. Make as many different pictures or objects as you can, and put as many ideas as you can into each one. Make them tell as complete and as interesting a story as you can. Add names or titles in the spaces provided. You will have ten minutes, so work quickly and carefully." (Thirty stimuli are provided.)

Wallach-Kogan procedures (3).

1. "Instances" (I). "This is a test to find out how many things you can think of which are in the same general group. I am going to tell you something and it will be your job to name as many things as you can that are like what I tell you. For example, I might say, 'things that are soft to touch.' Now you name some things that are soft to touch." (Let the child try.) "Yes, those are good answers." (If an answer is inappropriate, correct S.) "Some of the other things are a kitten, snow, a breeze, ice cream, or clothes." (Give these as though they are correct answers. Vary so as to include those not provided by S.) "So you see this test has many different right answers. Do you understand what you are to do? Now remember, I will name something and you name as many things as you can think of that are like what I've said. Okay, let's begin." (If S uses cues in the room, say, "Just close your eyes and name things that . . .")

2. "Alternate Uses" (AU). "This test is to find out how well you can think of different ways to use ordinary things. I will name an object--any kind of object, like a lightbulb or the floor--and it will be your job to tell me all the different ways you can think of that this object could be used. Any object can be used in a lot of different ways. For example, think of string. What are some of the ways you can think of that you could use string?" (Let the child try.) "Yes, those are right." (If an answer is inappropriate, correct S.) "Other right answers would be to attach to a fish hook, to jump rope, to sew with, to hang clothes on, and to pull shades." (Vary answers so as not to duplicate any the child has provided.) "So you see, there are many correct answers in this test. Do you understand what you are to do? Okay, let's begin then." (Precede each item with, "Tell me all the different ways you could use a ______.")

3. "Line Meanings" (LM). "This next test will help me find out how many different ways you can see things. I am going to show you some lines, and after you have looked at each one, I want you to tell me all the things it might be. Be sure that when you look at the line you tell me what the whole line makes you think of, not just part of it." (Present the first line.) "Here is the first line. You can turn it any way that you want to. Tell me all the things you can about it. What does it make you think of?" (Introduce each succeeding line by saying, "Here is the next one. Tell me all the things it might be. Remember, you can turn it any way you like.")

NEV Condition

Torrance procedures (5).

1. "Ask and Guess--Asking" (AGA). "Let's try a couple of games using this picture I have here." (Hand picture to S.) "The idea of this first game is to see how many questions you can think of which will help you to find out the things you don't know about the picture just from looking at it. Now, look at the picture. Ask yourself: 'What might be happening?' 'What can I tell for sure?' 'What else do I need to know to understand what is happening?' Don't worry about how good they are; just tell me all the different questions that come to your mind which can't be answered just by looking at the picture, but would help you to know more about it. Do you understand? Then let's begin."

2. "Ask and Guess--Guessing Causes" (AGG). "In this game, we'll use the same picture, but instead of asking questions, this will be a guessing game. Your part is to look at the picture again, and try to guess all the possible causes you can think of for the action shown in the picture. In other words, what kinds of things could have happened before to lead up to the scene in the picture? You may use things that might have happened just before the things happening in the picture, or something that happened a long time ago that made these things happen. Don't worry if any of your guesses seem silly; just make as many as you can. Do you understand? Then here we go."

3. "Unusual Uses" (UU). "Here's a(nother) game where you can really let your imagination run free. Most people throw

their empty cardboard boxes away, but these boxes have thousands of interesting and unusual uses. Suppose you were an inventor and you were trying to imagine as many of these interesting and unusual uses as possible. Pretend that you could use any size cardboard box at all, and that you had as many boxes as you wanted. Tell me all of the new and different uses that come into your mind, no matter what they are. Do you understand? Then go ahead."

4. "Just Suppose" (JS). "Here's a pretend game called 'Just Suppose' that should be lots of fun. I'm going to give you an improbable situation--it will probably never happen, but let's pretend it does. In your imagination, _just suppose_ this strange thing were to happen. Then try to imagine all the other interesting and exciting things that might happen because of it. _Okay, just suppose that clouds had strings attached to them which hung down to earth._ Now, imagine all of the other things that might happen because of this. How might it make things different? Make as many guesses as you want to."

5. "Repeated Figures" (RF). "Here is a drawing game. Just for fun, let's see how many different objects or pictures you can make from these pairs of straight lines I have here. You can place marks between lines, on the lines and outside the lines--wherever you want, in order to make your pictures. Make anything you like, but try to make the pairs of straight lines the main part of your drawing. Try to put lots of ideas into each one, and make whatever _you_ think will be most interesting and unusual. Since this is just for fun, don't worry about how

good your drawings look. When you have finished each one, make up a title to go with it, and write it on the line at the bottom. You may use this pencil and also these crayons if you want to." (Thirty stimuli are provided. This task is timed inconspicuously with a wristwatch to provide a fluency score--time: ten minutes.)

Wallach-Kogan procedures (3).

1. "Instances" (I). "In this game, I am going to tell you something and your part will be to name as many things as you can think of that are like what I tell you. For example, I might say, 'things that are soft to touch.' Now you name some things that are soft to touch." (Let the S try.) "Yes, those are fine. Some of the other things might be a kitten, snow, a breeze, ice cream, or clothes." (Give these in a suggesting manner. Vary suggestions to include those not provided by the child.) "So we see that there are all kinds of different answers in this game. Do you see how we play?" (If a child already indicates strong understanding, substitute, "I can see that you already know how we play this one.") "Now, again, I will name something and you name as many things as you can think of that are like what I have said. Okay, let's go." (If the child appears to be using cues in the testing room, say, "Just close your eyes and try to imagine all the things that . . .")

2. "Alternate Uses" (AU). "Now, in this game, I am going to name an object--any kind of object, like a lightbulb or the floor--and your part will be to tell me lots of different ways that the object could be used. Any object can be used in a lot of different ways. For example, think about string. What

are some of the ways you can think of that you might use string?" (Let the child try.) "Yes, those are fine. Some other things that occurred to me were that you could also use string to attach a fish hook, to jump rope, to sew with, to hang clothes on, and to pull shades." (Vary suggestions so as not to duplicate any the child has provided.) "There are lots more too, and yours are very good examples. I can see that you already understand how we play this game. So let's begin now. And remember, think of all the different ways you could use the object I name. Here we go." (Precede each item with, "Tell me the different ways you could use a _____.")

3. "Line Meanings" (LM). "This game is called the 'Line Game.' I am going to show you some lines, and after you have looked at each one, tell me all the things it makes you think of. Now take your time, and be sure that when you look at the line you tell me what the whole line makes you think of, and not just part of it. Okay?" (Present first line.) "Here is the first line. You can turn it any way you want to. Tell me all the things you can about it. What does it make you think of?" (Introduce each succeeding line by saying, "Here is the next one. Tell me all the things it makes you think of. Remember, you can turn it any way you like.")

After all eight tasks are completed, S is asked for his opinion as to which ones he liked best, least, and why. He also is asked for any other reactions he may have. This can be quite informal, but comments and responses to "liked best," "liked least" inquiries are recorded.

Appendix C

SCORING GUIDE FOR JUDGING "SAMENESS-DIFFERENCE" IN WALLACH-KOGAN RESPONSES

1. Responses qualified by adjectives and other modifiers which alter the meaning of the response in any way should be scored as different.

2. Responses containing nonessential descriptive adjectives should be considered the same as the identical response without such adjectives.

3. When a response appears incompletely specified, score according to the most common usage of that response.

4. If two responses can be conceptually differentiated into superordinate and subordinate concepts, (e.g., "Ball;" "Baseball"), they should be considered different.

5. In the "Line Meanings" subtasks, all conceivable responses should be considered appropriate, whether determined by part or all of the stimulus line. The lines may be turned in any direction.

6. Responses indicating "part of" an object are the same as the object simply named, unless specific parts are named; e.g., "Book" and "Edge of book" are identical; "Page of book" is not.

7. Other guidelines are left to the judges' discretion.

APPENDIX D

TABLES

Table 1. Guilford's Divergent Thinking Factors

(See pp. 5-6.)

Table 2. Comparison of EV and NEV Treatment Groups on Five Selection Variables

Variable	EV(N=40)		NEV(N=40)		F^a	p
	Mean	S.D.	Mean	S.D.		
Age, in months	142.70	7.30	142.15	5.65	0.14	n.s.
Social Class	3.30	0.94	3.02	1.12	1.42	n.s.
Verbal IQ	106.72	14.39	109.38	13.29	0.73	n.s.
Nonverbal IQ	105.10	11.37	103.28	14.23	0.40	n.s.
Total IQ	106.25	12.07	106.60	12.61	0.02	n.s.

a. For 1, 38 df, F's of 4.08 and 7.31 are significant at the .05 and .01 level, respectively.

Table 3. Intercorrelations among the Nine Wallach-Kogan Creativity Measures and with Intelligence for EV (N=40)

Variable	2	3	4	5	6	7	8	9	VIQ	NVIQ	TIQ
1. I-Number	.38[a]	.68	.80	.51	.75	.85	.50	.67	.08	.22	.15
2. AU-Number		.30	.32	.64	.20	.48	.81	.28	.04	-.01	.02
3. LM-Number			.53	.07	.82	.54	.24	.95	-.02	.05	.01
4. I-Originality				.52	.62	.65	.42	.53	.21	.38	.31
5. AU-Originality					.19	.43	.75	.07	.27	.13	.22
6. LM-Originality						.58	.23	.91	-.03	-.01	-.02
7. I-Flexibility							.54	.54	-.03	.14	.04
8. AU-Flexibility								.27	.24	.19	.23
9. LM-Flexibility									-.04	-.01	-.02

[a] For 38 df, r's of .31 and .40 are significant at the .05 and .01 levels, respectively. (These levels apply to all subsequent tables unless otherwise indicated.)

Table 4. Intercorrelations among the Twelve Torrance Verbal Creativity Measures and with Intelligence for EV (N=40)

Variable	2	3	4	5	6	7	8	9	10	11	12	VIQ	NVIQ	TIQ
1. AGA-Num	.71	.26	.30	.59	.27	.09	-.03	.75	.73	.26	-.27	.03	.20	.11
2. AGG-Num		.40	.46	.49	.16	.28	-.16	.59	.72	.43	-.23	.12	.22	.18
3. UU-Num			.40	.07	.22	.23	.37	.21	.40	.50	.29	.16	.22	.20
4. JS-Num				.13	.30	.36	-.02	.10	.40	.49	.16	.11	.26	.19
5. AGA-Orig					.56	.35	.06	.76	.58	.20	-.20	.18	.28	.24
6. AGG-Orig						.14	.25	.43	.57	.03	.01	.06	.16	.12
7. UU-Orig							.25	.15	.21	.65	-.08	.36	.20	.31
8. JS-Orig								.02	.20	.11	.38	.07	.08	.08
9. AGA-Flex									.75	.22	-.29	.11	.22	.27
10. AGG-Flex										.38	.19	.11	.25	.18
11. UU-Flex											-.03	.30	.27	.31
12. JS-Flex												-.20	-.02	-.13

Table 5. Intercorrelations between the Torrance Verbal and Wallach-Kogan Creativity Measures for EV (N=40)

		Torrance Measures											
	Variable	AGA-Num	AGG-Num	UU-Num	JS-Num	AGA-Orig	AGG-Orig	UU-Orig	JS-Orig	AGA-Flex	AGG-Flex	UU-Flex	JS-Flex
Wallach-Kogan Measures	I-Num	.39	.42	.57	.59	.26	.44	.33	.21	.29	.51	.45	.20
	AU-Num	-.03	.12	.60	.39	-.09	.18	.39	.28	.02	.19	.60	.17
	LM-Num	.15	.26	.40	.35	.13	.32	.14	.07	.15	.32	.18	.25
	I-Orig	.30	.38	.44	.73	.31	.46	.50	-.05	.22	.43	.46	.06
	AU-Orig	.09	.13	.46	.46	-.02	.18	.60	.23	.07	.28	.62	-.02
	LM-Orig	.19	.28	.33	.42	.16	.33	.28	-.05	.12	.31	.22	.13
	I-Flex	.22	.36	.56	.68	.03	.39	.11	.26	.13	.40	.39	.32
	AU-Flex	.13	.27	.59	.44	.02	.19	.46	.25	.14	.40	.78	.03
	LM-Flex	.20	.30	.41	.36	.17	.28	.18	-.01	.19	.32	.25	.20

Table 6. Intercorrelations between the Creativity Indexes, Intelligence, and Total Number, Originality, and Flexibility Scores for EV (N=40)

Comparison	r
Avg. r among Wallach-Kogan Totals (number, originality, flexibility)	.88
Avg. r among Torrance Verbal Totals (number, originality, flexibility)	.63
Wallach-Kogan Index *vs.* Torrance Index	.60
Wallach-Kogan Index *vs.* Total IQ	.12
Torrance Index *vs.* Total IQ	.30

Table 7. Intercorrelations among the Nine Wallach-Kogan Creativity Measures and with Intelligence for NEV (N=40)

Variable	2	3	4	5	6	7	8	9	VIQ	NVIQ	TIQ
1. I-Num	.50	.47	.68	.56	.54	.83	.60	.56	.08	.26	.19
2. AU-Num		.55	.40	.46	.36	.43	.68	.48	-.10	.01	-.04
3. LM-Num			.36	.47	.60	.40	.60	.86	.08	.28	.19
4. I-Orig				.78	.80	.59	.54	.63	-.16	.08	-.03
5. AU-Orig					.86	.42	.71	.71	-.17	.06	-.05
6. LM-Orig						.49	.67	.86	-.04	.23	.11
7. I-Flex							.64	.52	.18	.41	.33
8. AU-Flex								.72	-.01	.30	.17
9. LM-Flex									.09	.38	.26

Table 8. Intercorrelations among the Twelve Torrance Verbal Creativity Measures and with Intelligence for NEV (N=40)

Variable	2	3	4	5	6	7	8	9	10	11	12	VIQ	NVIQ	TIQ
1. AGA-Num	.66	.57	.64	.50	.13	.21	-.01	.77	.31	.25	.07	-.02	.23	.12
2. AGG-Num		.38	.64	.44	.30	-.04	.02	.65	.62	.00	.16	.01	.18	.10
3. UU-Num			.41	.29	.10	.13	.18	.39	.01	.54	-.03	-.08	.11	.03
4. JS-Num				.26	.13	.05	-.07	.47	.30	.07	.15	.17	.14	.17
5. AGA-Orig					.22	.30	.13	.67	.51	.27	.06	-.18	.12	-.02
6. AGG-Orig						.18	.06	.20	.61	.30	.36	-.27	-.22	-.26
7. UU-Orig							.28	.04	.06	.73	.16	-.39	-.12	-.27
8. JS-Orig								.05	.09	.34	.58	-.03	.08	.03
9. AGA-Flex									.53	.14	.09	-.19	.09	-.05
10. AGG-Flex										.07	.31	-.21	.05	-.08
11. UU-Flex											.24	-.28	-.01	-.15
12. JS-Flex												.04	.03	.03

Table 9. Intercorrelations between the Torrance Verbal and Wallach-Kogan Creativity Measures for NEV (N=40)

		Torrance Measures											
	Variable	AGA-Num	AGG-Num	UU-Num	JS-Num	AGA-Orig	AGG-Orig	UU-Orig	JS-Orig	AGA-Flex	AGG-Flex	UU-Flex	JS-Flex
Wallach-Kogan Measures	I-Num	.50	.57	.40	.44	.19	.26	.05	.41	.47	.33	.18	.39
	AU-Num	48	.36	.68	.25	.14	.30	.12	.23	.44	.19	.54	.17
	LM-Num	.20	.23	.42	.12	.07	.18	.07	.49	.12	.30	.32	.33
	I-Orig	.25	.35	.37	.15	.26	.42	.04	.13	.42	.33	.22	.08
	AU-Orig	.30	.44	.40	.07	.23	.39	.12	.15	.31	.29	.28	.14
	LM-Orig	.12	.24	.38	-.05	.23	.35	.12	.33	.15	.23	.28	.17
	I-Flex	.29	.32	.28	.23	.13	.23	.14	.43	.25	.19	.27	.27
	AU-Flex	.38	.26	.42	.06	.16	.30	.35	.34	.23	.21	.50	.20
	LM-Flex	.20	.27	.40	.01	.18	.32	.10	.42	.18	.36	.34	.31

Table 10. Summary of Creativity-Intelligence Relationships for All Number, Originality, and Flexibility Variables for EV (N=40)

Variables Compared	# of r's	Avg.r	Range	%sig. p<.05	%sig. p<.01
Wallach-Kogan Measures	36	.52	.07to.95	90	70
Torrance Measures	64	.28	-.29to.76	42	34
Wallach-Kogan *vs.* Torrance	108	.29	-.09to.78	43	26
Wallach-Kogan *vs.* IQ	27	.10	-.04to.38	07	00
Torrance *vs.* IQ	36	.16	-.20to.36	09	03

Table 11. Summary of Creativity-Intelligence Relationships for All Number, Originality, and Flexibility Variables for *NEV* (N=40)

Variables Compared	# of r's	Avg. r	Range	%sig. $p < .05$	%sig. $p < .01$
Wallach-Kogan Measures	36	.59	.36to.86	100	92
Torrance Measures	64	.28	-.07to.77	36	28
Wallach-Kogan *vs.* Torrance	108	.27	-.05to.68	39	19
Wallach-Kogan *vs.* IQ	27	.11	-.17to.41	11	04
Torrance *vs.* IQ	36	-.03	-.39to.17	03	00

Table 12. Intercorrelations between the Creativity Indexes, Intelligence, and Total Number, Originality, and Flexibility Scores for *NEV* (N=40)

Comparison	r
Avg. r among Wallach-Kogan Totals (number, originality, flexibility)	.76
Avg. r among Torrance Verbal Totals (number, originality, flexibility)	.55
Wallach-Kogan Index *vs.* Torrance Index	.49
Wallach-Kogan Index *vs.* Total IQ	.14
Torrance Index *vs.* Total IQ	-.06

Table 13. Intercorrelations between "Repeated Figures" Variables and Other Creativity and Intelligence Variables for EV and NEV Conditions

		EV				NEV			
	Variable	RF-Num	RF-Orig	RF-Flex	RF-Elab	RF-Num	RF-Orig	RF-Flex	RF-Elab
Wallach-Kogan Measures	I-Num	-.31	.06	-.18	.24	.24	-.05	-.02	.23
	AU-Num	-.06	.14	.17	.15	.28	-.08	.00	.18
	LM-Num	-.17	.24	.14	.23	.21	-.02	.11	.32
	I-Orig	-.21	.04	-.17	.13	.23	.13	.20	.05
	AU-Orig	-.19	.04	-.10	.14	.17	.23	.30	.20
	LM-Orig	-.34	.23	-.07	.17	.12	.18	.19	.25
	I-Flex	-.31	.03	-.19	.13	.13	.00	.08	.15
	AU-Flex	-.30	.09	-.02	.18	.10	.15	.19	.31
	LM-Flex	-.22	.29	.10	.24	.14	.08	.15	.36
Torrance Measures	AGA-Num	-.08	-.05	-.11	.11	.18	-.06	.02	.02
	AGG-Num	-.33	.19	-.03	.30	.28	.02	.19	.10
	UU-Num	-.13	.30	.15	.28	.21	-.08	-.06	.05
	JS-Num	-.27	.08	-.16	.08	.40	-.28	-.07	.08
	AGA-Orig	-.15	.05	.05	.15	-.10	.22	.16	-.04
	AGG-Orig	.00	-.20	-.15	-.05	.01	.16	.19	-.04
	UU-Orig	-.29	.27	.11	.20	.06	.21	.22	.06
	JS-Orig	-.05	-.04	.01	-.09	-.14	.03	-.10	.06
	AGA-Flex	-.12	.11	.12	.32	.22	-.04	.02	-.04
	AGG-Flex	-.26	-.07	-.14	.11	-.04	.15	.14	.08
	UU-Flex	-.31	.25	.07	.34	.16	.01	.08	.04
	JS-Flex	-.12	.14	-.02	-.22	.07	-.17	-.26	.04
IQ Measures	VIQ	.19[a]	.36	.26	.30	-.18	-.18	-.15	.22
	NVIQ	.06	.23	.13	.18	-.18	-.18	-.09	.18
	TIQ	.14	.33	.21	.26	-.19	-.20	-.13	.22

[a] The uncorrected fluency score is used in correlations with IQ, as there appeared to be substantial difference between fluency and number correlations with IQ.

Table 14. Results of One-way Analyses of Variance across Assessment Condition for All Measures of All Procedures, and Totals

Variable	EV Mean	EV S.D.	NEV Mean	NEV S.D.	Diff.	F	p[a]
Wallach-Kogan Tasks							
I-Flu	52.58	26.80	61.18	29.78	8.60	1.84	n.s.
I-Orig	6.92	7.94	9.45	11.46	2.53	1.31	n.s.
I-Flex	23.95	6.38	26.58	6.84	2.63	3.15	.10
AU-Flu	22.50	12.55	25.43	11.76	2.93	1.16	n.s.
AU-Orig	5.10	7.69	6.05	7.49	0.95	0.31	n.s.
AU-Flex	14.85	5.49	16.18	5.12	1.33	1.25	n.s.
LM-Flu	52.68	29.46	64.65	37.92	11.97	2.49	n.s.
LM-Orig	14.40	15.10	20.70	24.48	6.30	1.92	n.s.
LM-Flex	38.92	17.68	45.02	18.86	6.10	2.23	n.s.
Torrance Tasks							
AGA-Flu	13.82	14.25	14.90	12.77	1.08	0.13	n.s.
AGA-Orig	5.22	4.57	5.98	5.42	0.76	0.45	n.s.
AGA-Flex	6.77	3.60	7.70	3.51	0.93	1.35	n.s.
AGG-Flu	8.90	5.38	9.72	6.11	0.82	0.41	n.s.
AGG-Orig	3.90	2.57	4.30	3.10	0.40	0.40	n.s.
AGG-Flex	5.20	2.20	5.28	2.35	0.08	0.02	n.s.
UU-Flu	13.58	10.30	12.50	7.58	-1.08	0.28	n.s.
UU-Orig	5.75	4.62	5.92	5.06	0.13	0.03	n.s.
UU-Flex	6.55	3.37	6.75	2.92	0.20	0.08	n.s.
JS-Flu	9.52	6.11	10.32	6.19	0.80	0.34	n.s.
JS-Orig	2.30	2.03	3.52	3.44	1.22	3.76	.10
JS-Flex	3.05	2.70	3.90	3.45	0.85	1.51	n.s.
RF-Flu	12.45	5.38	10.20	3.94	-2.25	4.56	.05
RF-Orig	11.28	5.86	9.30	4.27	-1.98	2.96	.10
RF-Flex	9.32	3.64	7.82	3.32	-1.50	3.71	.10
RF-Elab	28.80	9.61	29.00	14.18	0.20	0.01	n.s.
Wallach-Kogan Totals							
Flu	127.75	58.42	151.25	71.59	25.50	2.59	n.s.
Orig	26.42	24.65	36.20	41.00	9.78	1.67	n.s.
Flex	77.72	24.39	87.77	27.32	10.05	3.01	n.s.
Torrance Verbal Totals							
Flu	45.82	27.77	47.45	25.95	1.63	0.07	n.s.
Orig	17.18	9.50	19.72	11.03	2.54	1.23	n.s.
Flex	21.58	7.11	23.62	7.94	2.04	1.48	n.s.
Torrance Total Totals							
Flu	58.28	28.80	57.65	26.83	-0.63	0.01	n.s.
Orig	28.45	11.62	29.03	12.82	0.58	0.04	n.s.
Flex	30.90	8.11	31.45	8.51	0.55	0.09	n.s.

a. For 1,78 df, F's of 2.79, 3.96, and 6.96 are significant at p=.10, .05, and .01, respectively.

Table 15. Summary of Creativity-Intelligence Relationships for Number and Originality Variables for EV (N=40)

Variables Compared	# of r's	Avg.r	Range	%sig. $p<.05$	%sig. $p<.01$
Wallach-Kogan Measures	15	.49	.07to.72	73	60
Torrance Measures	28	.27	-.16to.71	35	25
Wallach-Kogan *vs.* Torrance	48	.29	-.09to.60	50	29
Wallach-Kogan *vs.* IQ	18	.11	-.03to.38	11	00
Torrance *vs.* IQ	24	.17	.03to.36	08	04

Table 16. Summary of Creativity-Intelligence Relationships for Flexibility Variables for EV (N=40)

Variables Compared	# of r's	Avg.r	Range	%sig. $p<.05$	%sig. $p<.01$
Wallach-Kogan Measures	21	.54	.07to.95	76	76
Torrance Measures	38	.27	-.29to.76	45	40
Wallach-Kogan *vs.* Torrance	60	.32	-.02to.78	37	25
Wallach-Kogan *vs.* IQ	9	.08	-.04to.23	00	00
Torrance *vs.* IQ	12	.14	-.20to.31	11	00

Table 17. Summary of Creativity-Intelligence Relationships for Number and Originality Variables for NEV (N=40)

Variables Compared	# of r's	Avg.r	Range	%sig. $p<.05$	%sig. $p<.01$
Wallach-Kogan Measures	15	.56	.36to.86	100	87
Torrance Measures	28	.25	-.07to.66	29	25
Wallach-Kogan *vs.* Torrance	48	.26	-.05to.68	40	25
Wallach-Kogan *vs.* IQ	18	.05	-.17to.28	00	00
Torrance *vs.* IQ	24	-.01	-.39to.23	04	00

Table 18. Summary of Creativity-Intelligence Relationships for Flexibility Variables for *NEV* (N=40)

Variables Compared	# of r's	Avg. r	Range	%sig. $p<.05$	%sig. $p<.01$
Wallach-Kogan Measures	21	.62	.40 to .86	100	100
Torrance Measures	38	.37	-.03 to .77	33	29
Wallach-Kogan *vs.* Torrance	60	.27	.01 to .54	38	15
Wallach-Kogan *vs.* IQ	9	.23	-.01 to .41	33	11
Torrance *vs.* IQ	12	-.06	-.28 to .09	00	00

Table 19. Intercorrelations between Modified Creativity Indexes, Intelligence, and Total Number and Originality Scores for *EV* and *NEV*

Comparison	r(EV)	r(NEV)
Avg. r among Wallach-Kogan Totals (number and originality only)	.87	.67
Avg. r among Torrance Verbal Totals (number and originality only)	.46	.35
Wallach-Kogan Index *vs.* Torrance Index	.60	.47
Wallach-Kogan Index *vs.* IQ	.14	.10
Torrance Index *vs.* IQ	.31	-.05

Table 20. Summary of Creativity-Intelligence Relationships by Assessment Condition and Scoring Method

	Variables compared	Avg. r Num, Orig	Avg. r Num, Orig, Flex
EV	Wallach-Kogan Measures	.49	.52
	Torrance Measures	.27	.28
	Wallach-Kogan *vs.* Torrance	.29	.29
	Wallach-Kogan *vs.* IQ	.11	.10
	Torrance *vs.* IQ	.17	.16
NEV	Wallach-Kogan Measures	.56	.59
	Torrance Measures	.25	.28
	Wallach-Kogan *vs.* Torrance	.26	.27
	Wallach-Kogan *vs.* IQ	.05	.11
	Torrance *vs.* IQ	-.01	-.03

Table 21. Results of Two-way Analyses of Variance between Social Class and Assessment Condition for All Measures of All Procedures

	MC	LC				
Variable	Mean	Mean	F(SES)	p[a]	F(int)	p[a]
Wallach-Kogan Tasks						
I-Flu	67.5	54.1	1.87	n.s.	0.13	n.s.
I-Orig	13.1	7.1	2.66	n.s.	0.57	n.s.
I-Flex	28.2	24.8	2.06	n.s.	2.10	n.s.
AU-Flu	28.0	22.5	1.92	n.s.	0.37	n.s.
AU-Orig	9.0	4.3	4.08	.05	0.79	n.s.
AU-Flex	16.9	15.1	0.64	n.s.	0.20	n.s.
LM-Flu	67.9	56.8	0.43	n.s.	0.22	n.s.
LM-Orig	25.0	15.6	1.24	n.s.	0.36	n.s.
LM-Flex	45.9	41.6	0.14	n.s.	0.39	n.s.
Torrance Tasks						
AGA-Flu	14.0	13.6	0.61	n.s.	0.99	n.s.
AGA-Orig	5.8	5.7	0.04	n.s.	0.79	n.s.
AGA-Flex	7.1	7.2	0.18	n.s.	0.38	n.s.
AGG-Flu	8.3	9.8	1.35	n.s.	0.02	n.s.
AGG-Orig	3.6	5.0	2.69	n.s.	0.08	n.s.
AGG-Flex	4.9	5.3	0.50	n.s.	0.02	n.s.
UU-Flu	13.8	12.3	0.51	n.s.	0.61	n.s.
UU-Orig	6.0	5.6	0.22	n.s.	1.99	n.s.
UU-Flex	10.8	6.6	0.48	n.s.	0.34	n.s.
JS-Flu	10.9	10.4	0.13	n.s.	0.61	n.s.
JS-Orig	3.2	3.0	0.01	n.s.	0.14	n.s.
JS-Flex	2.8	4.1	2.75	n.s.	1.98	n.s.
RF-Flu	11.2	10.6	0.62	n.s.	0.41	n.s.
RF-Orig	10.8	9.6	1.72	n.s.	0.30	n.s.
RF-Flex	8.7	8.3	0.18	n.s.	0.09	n.s.
RF-Elab	32.0	28.3	0.72	n.s.	0.28	n.s.

a. For 1, 52 df, F's of 2.80, 4.03, and 7.31 are significant at p=.10, .05, and .01, respectively.

Table 22. Results of Two-way Analyses of Variance between Intelligence and Assessment Condition for All Measures of All Procedures

	Hi IQ	Lo IQ				
Variable	Mean	Mean	F(IQ)	p^a	F(int)	p^a
		Wallach-Kogan Tasks				
I-Flu	58.7	55.1	0.31	n.s.	1.07	n.s.
I-Orig	8.2	8.1	0.00	n.s.	3.02	.10
I-Flex	26.4	24.2	2.24	n.s.	0.12	n.s.
AU-Flu	24.0	23.8	0.01	n.s.	1.70	n.s.
AU-Orig	5.8	5.3	0.07	n.s.	2.08	n.s.
AU-Flex	16.4	14.7	2.01	n.s.	0.03	n.s.
LM-Flu	59.1	58.3	0.01	n.s.	0.10	n.s.
LM-Orig	17.3	17.8	0.01	n.s.	0.00	n.s.
LM-Flex	42.9	41.2	0.18	n.s.	0.60	n.s.
		Torrance Tasks				
AGA-Flu	15.2	13.4	0.36	n.s.	2.06	n.s.
AGA-Orig	5.6	5.5	0.02	n.s.	2.91	.10
AGA-Flex	7.1	7.3	0.03	n.s.	3.10	.10
AGG-Flu	8.9	9.6	0.26	n.s.	5.48	.05
AGG-Orig	3.6	4.5	2.38	n.s.	7.26	.01
AGG-Flex	5.1	5.3	0.18	n.s.	4.96	.05
UU-Flu	14.2	11.8	1.55	n.s.	5.01	.05
UU-Orig	5.9	5.7	0.04	n.s.	2.73	n.s.
UU-Flex	6.8	6.4	0.30	n.s.	2.51	n.s.
JS-Flu	10.6	9.2	1.06	n.s.	1.25	n.s.
JS-Orig	3.2	2.6	0.75	n.s.	0.50	n.s.
JS-Flex	3.3	3.6	0.13	n.s.	0.26	n.s.
RF-Flu	10.9	11.6	0.43	n.s.	1.07	n.s.
RF-Orig	10.4	10.1	0.12	n.s.	3.63	.10
RF-Flex	8.6	8.6	0.00	n.s.	1.20	n.s.
RF-Elab	30.9	27.0	2.10	n.s.	0.80	n.s.

a. For 1, 78 df, F's of 2.79, 3.96, and 6.96 are significant at p=.10, .05, and .01, respectively.

REFERENCES

REFERENCES

Allport, G.W. *Pattern and growth in personality.* New York: Holt, Rinehart, and Winston, 1961, p. 229.

Anderson, H.H. (Ed.). *Creativity and its cultivation.* New York: Harper, 1959.

Anderson, H.H. On the meaning of creativity. In H.H. Anderson (Ed.), *Creativity in childhood and adolescence.* Palo Alto: Science and Behavior Books, Inc., 1965, 46-61.

Bereiter, C. Fluency abilities in preschool children. *J. genet. Psychol.*, 1961, *98*, 47-48.

Bereiter, C., and Englemann, S. *Teaching disadvantaged children in the preschool.* Englewood Cliffs, N.J.: Prentice Hall, 1966.

Circirelli, V.G. The relationship between measures of creativity, IQ, and academic achievement; interaction and threshold effects. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, 1964.

Clark, C.M., Veldman, D.J., and Thorpe, J.S. Convergent and divergent thinking abilities of talented adolescents. *J. educ. Psychol.*, 1965, *56*, 157-163.

Cline, V.B., Richards, J.M., and Needham, W.E. Creativity tests and achievement in high school science. *J. appl. Psychol.*, 1963, *47*, 184-189.

DeMille, R., and Merrifield, P.R. Getzels and Jackson's "Creativity and intelligence: explorations with gifted students." *Educ. psychol. Measmt.*, 1962, *22*, 803-808.

Dentler, R.A., and Mackler, B. Originality: some social and personal determinants. *Behav. Sci.*, 1964, *9*, 1-7.

Feldhusen, J.F., Denny, T., and Condon, C.F. Anxiety, divergent thinking, and achievement. *J. educ. Psychol.*, 1965, *56*, 40-45.

Flescher, I. Anxiety and achievement of intellectually gifted and creatively gifted children. *J. Psychol.*, 1963, *56*, 251-268.

Getzels, J.W., and Jackson, P.W. Creativity and intelligence. New York: Wiley, 1962.

Gewirtz, J.L. Studies in word fluency: I. Its relation to vocabulary and mental age in young children. J. genet. Psychol., 1948, 72, 165-176.

Ghiselin, B. (Ed.). The creative process. Berkeley: University of California Press, 1952.

Guilford, J.P. The structure of intellect. Psychol. Bull., 1956, 53, 267-293.

Guilford, J.P. Creative abilities in the arts. Psychol. Rev., 1957, 64, 110-118.

Guilford, J.P. Three faces of intellect. Amer. Psychol., 1959, 14, 469-479.

Guilford, J.P. A psychometric approach to creativity. In H.H. Anderson (Ed.), Creativity in childhood and adolescence. Palo Alto: Science and Behavior Books, 1965, 1-19.

Guilford, J.P. Intelligence: 1965 model. Amer. Psychol., 1966, 21, 20-26.

Guilford, J.P., and Christensen, P.R. A factor analytic study of verbal fluency. Rep. psychol. Lab., 1956, #17, Los Angeles: University of Southern California.

Guilford, J.P., Frick, J.W., Christensen, P.R., and Merrifield, P.R. A factor analytic study of flexibility in thinking. Rep. psychol. Lab., 1957, #18, Los Angeles: University of Southern California.

Horton, D.L., Marlowe, D., and Crowne, D.P. The effect of instructional set and need for social approval on commonality of word association responses. J. abn. soc. Psychol., 1963, 66, 67-72.

Lorge, I., and Thorndike, R.L. The Lorge-Thorndike intelligence tests, examiner's manual. Boston: Houghton Mifflin, 1957.

Mackler, B. Creativity and Life Style. Unpublished Doctoral dissertation, University of Kansas, Lawrence, 1962.

Madans, A.B. What is creativity? Trans. N.Y. acad. Sci., 1964, 26, 781-787.

Maslow, A.H. Creativity in self-actualizing people. In H.H. Anderson (Ed.), Creativity and its cultivation. New York: Harper, 1959, 83-95.

Mednick, S.A. The associative basis of the creative process. Psychol. Rev., 1962, 69, 220-232.

Orpet, R.E., and Meyers, C.E. Six structure-of-intellect hypotheses in six-year-old children. J. educ. Psychol., 1966, 57, 341-346.

Reece, M.M., and Whitman, R.N. Expressive movements, warmth, and verbal reinforcement. J. abn. soc. Psychol., 1962, 64, 234-236.

Reid, J.B., King, F.J., and Wickwire, Pat. Cognitive and other personality characteristics of creative children. Psychol. Rep., 1959, 5, 729-738.

Ripple, R.E., and May, F.B. Caution in comparing creativity and IQ. Psychol. Rep., 1962, 10, 229-230.

Rogers, C.R. Toward a theory of creativity. In H.H. Anderson (Ed.), Creativity and its cultivation. New York: Harper, 1959, 69-82.

Rogers, C.R. Actualizing tendency in relation to "motives" and to consciousness. In M.R. Jones (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1963, 1-24.

Siipola, Elsa, Walker, W. Nannette, and Kolb, Dorothy. Task attitudes in association, projective and non-projective. J. Pers., 1955, 23, 441-459.

Smith, R.M. The relationship of creativity to social class. Pittsburgh: University of Pittsburgh Press, 1965. (Cooperative Research Report #2250).

Thorndike, R.L. Some methodological issues in the study of creativity. In Proceedings of the 1962 invitational conference on testing problems. Princeton: Educational Testing Service, 1963, 40-54.

Torrance, E.P. Guiding creative talent. Englewood Cliffs, N.J.: Prentice Hall, 1962.

Torrance, E.P. Education and the creative potential. Minneapolis: University of Minnesota Press, 1963.

Torrance, E.P. Education and creativity. In C.W. Taylor (Ed.), Creativity: progress and potential. New York: McGraw-Hill, 1964, 49-128.

Torrance, E.P. Torrance tests of creative thinking. Princeton: Personnel Press, 1966.

Wallach, M.A., and Kogan, N. *Modes of thinking in young children: a critical examination of the creativity-intelligence distinction.* New York: Holt, Rinehart, and Winston, 1965.

Wallach, M.A. Creativity. In P.H. Mussen (Ed.), *Manual of child psychology.* New York: John Wiley, in press.

Ward, W.C. Creativity and impulsivity in kindergarten children. Unpublished doctoral dissertation, Duke University, 1966.

Wild, Cynthia. Creativity and adaptive regression. *J. pers. soc. Psychol.*, 1965, *2*, 161-169.

Wilson, R.C., Guilford, J.P., Christensen, P.R., and Lewis, D.J. A factor analytic study of creative thinking abilities. *Psychometrika*, 1954, *19*, 297-311.

Wodtke, K.H. Some data on the reliability and validity of creativity tests at the elementary school level. *Educ. and psychol. Measmt.*, 1964, *24*, 399-408.

Yamamoto, K. Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. *Brit. j. educ. Psychol.*, 1965, *35*, 300-305.

Yamamoto, K., and Chimbidis, Maria E. Achievement, intelligence, and creative thinking in fifth grade children: A correlational study. *Merrill Palmer Quart.*, 1966, *12*, 233-241.

Zigler, E., and Butterfield, E.C. Motivational factors and IQ changes in culturally deprived children attending nursery school. *Child Devel.*, 1968, *39*, 1-14.

BIOGRAPHICAL NOTE

Name: David William Sherwood.

Date of birth: January 18, 1943.

Place of birth: Rochester, New York.

Universities attended and degrees: Wittenberg University, Springfield, Ohio, A.B. in Psychology, 1964; Duke University, Durham, North Carolina.

Publications: Sarbin, T.R., Wenk, E.A., and Sherwood, D.W., An effort to identify assault-prone offenders. *Journal of research in crime and Delinquency*, 1968, *5*, 66-71.

Positions held: U.S. Public Health Fellow, Duke University--September, 1964, to June, 1966; September, 1966, to June, 1967.

Student Professional Assistant in Psychology, Deuel Vocational Institution, Tracy, California--June, 1966, to August, 1966.

Intern in Clinical Psychology, Astor Home for Children, Rhinebeck, New York--July, 1967, to September, 1967; February, 1968, to the present.

Research Technician, Education Improvement Program, Durham, North Carolina--September, 1967, to January, 1968.

Honoraries: The Society of Sigma Xi.

Ph.D. 1968

Date in